LUISS

Corso di laurea in Giurisprudenza

Cattedra di Metodologia della Scienza Giuridica

La decisione robotica tra luci e ombre

Chiar. mo Prof. Antonio Punzi

RELATORE

Chiar. ma Prof. ssa Alessia Farano

CORRELATORE

Giuseppe Matteo Capaldo

matricola n. 156563

CANDIDATO

Anno Accademico 2024/2025

INTRODUZIONE

Il presente lavoro è dedicato allo studio dell'impatto prodotto dall'intelligenza artificiale in ambito giudiziario, con particolare riferimento alla possibilità di impiegare tale tecnologia per sostituire il giudice nella decisione dei casi concreti.

Preliminarmente, per delimitare l'ambito oggettivo del lavoro, è necessario comprendere il fenomeno dell'intelligenza artificiale (IA) cercando di ricostruire una nozione che appare molto complessa ed articolata.

Tradizionalmente l'IA è definita come un sistema informatico che agisce in un modo tale che, se ad agire fosse un essere umano, definiremmo "intelligente". Essa comprende sistemi che, attraverso l'acquisizione e l'interpretazione di dati, pensano e agiscono come esseri umani e, dunque, razionalmente, per raggiungere un determinato obiettivo.

È opportuno distinguere l'IA debole o ristretta (ANI) che è addestrata e orientata ad eseguire attività specifiche, dall' IA forte che comprende, sia l'IA generale (AGI) che conferisce ad una macchina un'intelligenza pari a quella umana e la capacità di risolvere problemi, imparare e pianificare il futuro, sia l'ASI o superintelligenza, che in futuro sarebbe destinata a superare l'intelligenza e le capacità del cervello umano.

Dopo aver sommariamente ricostruito la genesi e l'evoluzione della ricerca scientifica relativa all'IA, e aver individuato le caratteristiche, i pregi e i limiti dei relativi sistemi, si esamina il rapporto tra questi ultimi e il mondo del diritto. La tecnologia in esame, infatti, costituisce un utile strumento di ausilio per gli operatori giuridici. Allo stesso tempo, essa stessa deve essere regolamentata con norme non solo etiche, ma anche giuridiche.

Pertanto, il capitolo si conclude con l'analisi del Regolamento del Parlamento e del Consiglio 2024/1689 del 21.5.2024 (*AI Act*), che disciplina tutte le fasi di vita dell'IA con un approccio precauzionale, individuando i rischi ad essa connessi e provando ad evitarli o, almeno, attenuarli, con specifiche regole di condotta. Inoltre, si analizza la Convenzione quadro sull'intelligenza artificiale e i diritti umani del Consiglio d'Europa del 17.05.2024 e, infine, la strategia italiana per l'intelligenza artificiale 2024-2026 e il d.d.l. AS 1146.

Dopo aver chiarito cosa si debba intendere per IA e aver individuato l'impostazione di fondo che ispira le prime norme giuridiche che disciplinano il fenomeno, il secondo

capitolo del lavoro approfondisce il tema della giustizia predittiva. Essa è costituita dagli strumenti informatici basati su database tendenzialmente giurisprudenziali, che, con l'ausilio di algoritmi di ordinamento, permettono di anticipare l'orientamento giurisprudenziale. Si tratta di strumenti che consentono, in primo luogo, di automatizzare la ricerca in campo legale. Inoltre, nella realtà, vi sono sistemi che non si limitano ad individuare il tendenziale orientamento giurisprudenziale, ma che determinano anche le probabilità di soccombenza o vittoria delle parti, anticipando l'esito della singola controversia in termini statistici. Dopo aver ricostruito l'evoluzione dei sistemi di giustizia predittiva, si analizzano le caratteristiche, i benefici e i rischi connessi al loro uso, le esperienze e i progetti attuati in ambito internazionale e nazionale e, infine, si esamina la disciplina etica e giuridica dedicata a tale specifico tema.

L'ultimo capitolo è dedicato alla riflessione filosofica sull'utilizzo dei sistemi di IA a fini decisori.

Come, più in generale, per il rapporto tra uomo e macchina, anche per lo sviluppo e le potenzialità decisorie dell'IA, si sono delineati orientamenti di pensiero opposti. Da una parte, vi è chi vede nelle macchine una minaccia e, nell'IA, un demone pronto a prendere il sopravvento sulle persone e sull'umano. Dall'altra parte si tende a riconoscere alla macchina un ruolo salvifico fino ad aspirare e sognare un nuovo mondo dove macchine guidate da strumenti di IA amplifichino le capacità umane.

Dopo aver esaminato le opposte correnti di pensiero, si approfondisce la tesi del tecno umanesimo che postula la formazione di una cultura, al tempo stesso, tecnica ed umanistica che superi la separazione tra le tecnoscienze e il pensiero umanistico. Tale cultura conferisce alle macchine e all'IA un'equilibrata funzione di sostegno e integrazione dell'uomo.

I sistemi di IA, infatti, non devono spaventare e devono essere considerati come un mezzo di cui l'uomo non può e non deve diventare vittima. A tale scopo un ruolo fondamentale è svolto dal diritto attraverso il quale il fenomeno dell'IA deve essere disciplinato, mediante regole flessibili e dinamiche che si basano sui principi fondamentali dell'ordinamento.

CAPITOLO I

L'Intelligenza Artificiale

Sommario: 1.1 La nozione di *intelligenza artificiale*. -1.1.1 Le origini dell'intelligenza artificiale –1.1.2 Gli inverni dell'intelligenza artificiale e i sistemi esperti - 1.1.3 L'apprendimento automatico. - 1.2 L'intelligenza artificiale e il diritto. – 1.3 L'intelligenza artificiale nel diritto positivo: il Regolamento del Parlamento e del Consiglio 2024/1689 del 21.5.2024 (*AI Act*) e le Direttive sulla Responsabilità civile. – 1.4 La Convenzione quadro sull'intelligenza artificiale e i diritti umani del Consiglio d'Europa del 17.05.2024. -1.5 La strategia italiana per l'intelligenza artificiale 2024-2026 e il d.d.l. AS 1146.

1.1 La nozione di *intelligenza artificiale*

L'intelligenza artificiale è un fenomeno complesso nel quale confluiscono diverse tecnologie. Lo sviluppo di sistemi intelligenti è oggetto di studio dell'informatica e della matematica. Tuttavia, tali sistemi hanno un ambito di applicazione pressoché illimitato e, quindi, sono studiati anche dal punto di vista delle scienze biomediche, dell'economia, della sociologia, delle scienze cognitive, della filosofia e del diritto[1].

La complessità del fenomeno in esame non consente di elaborare una definizione univoca di "intelligenza artificiale"[2]. Come rileva un noto studioso[3], se chiedessimo a cento ricercatori di definire l'intelligenza artificiale, avremmo, probabilmente, almeno cento definizioni diverse.

I molteplici tentativi definitori possono essere classificati in base a diversi approcci[4].

L'approccio ontologico cerca di chiarire cosa sia l'intelligenza artificiale, confrontando il fenomeno con l'intelligenza umana[5]. Tuttavia, il concetto di "*intelligenza*", nelle

[1] L.C. AJELLO, M. DAPOR, *Intelligenza artificiale: i primi 50 anni*, in *Mondo digitale* 2004,2 , p.3; M.SOMALVICO, *Intelligenza artificiale*, HP, 1987, pp.1 e ss.

[2] Commissione Europea, Joint Research center, Technical Report "*AI Watch.Defining Artificial Intelligence. Towards an operational definition and taxonomy of artificial intelligence*" 2020.

[3] O. STOCK, *Intelligenza artificiale oggi e domani*, in *Giornale italiano di psicologia*,2018, 1, p.159.

[4] S. SAPIENZA, *Decisioni algoritmiche e diritto*, Giuffrè, Milano, 2024, p.11.

[5] Un esempio di approccio ontologico è costituito dalla definizione elaborata durante il seminario estivo di ricerca organizzato nel 1956 presso il Dartmouth College di Hanover nel New Hampshire.

scienze cognitive, è ambiguo[6]. Peraltro, applicare alle macchine, categorie concettuali tipiche del comportamento umano, è fuorviante e rende lacunose le definizioni di intelligenza artificiale.

L'approccio definitorio funzionale, invece, indaga su cosa sia in grado di fare l'intelligenza artificiale rispetto all'essere umano[7]. Secondo tale approccio afferiscono all'"*intelligenza artificiale*" i sistemi che pensano come esseri umani, i sistemi che pensano razionalmente, quelli che agiscono come esseri umani e quelli che agiscono razionalmente[8]. Da un punto di vista funzionale, si distingue[9], altresì, tra intelligenza artificiale debole (ANI)[10] che può svolgere compiti complessi con riferimento ad una sfera limitata di capacità, e intelligenza artificiale forte (AGI)[11] che, invece, sarebbe in grado di raggiungere stati cognitivi assimilabili agli stati mentali umani. Malgrado l'intelligenza artificiale forte sia, al momento, un'ipotesi teorica, si teme[12] che in futuro si possa giungere anche ad una super intelligenza artificiale (ASI) costituita da sistemi con funzioni superiori alle abilità cognitive umane.

Seguendo l'approccio eziologico[13] l'intelligenza artificiale è definita attraverso gli elementi comuni ai vari sistemi in essa sussumibili. L'High Level Expert Group on AI

L'intelligenza artificiale fu definita come un sistema informatico che agisce in modo tale che, se ad agire fosse un essere umano, definiremmo tale agire come intelligente.

[6] J. KAPLAN, *AI's PR Problem. Had artificial intelligence been named something less spooky, we'd probably worry about it less*, in *MIT Tech. Rev.*, 3, 2017, www.technologyreview.com. Dal punto di vista filosofico e scientifico, la definizione di cosa sia l'intelligenza è connessa alla questione irrisolta della individuazione della natura della coscienza. Peraltro, almeno per ora, si ritiene sia improprio associare il concetto di intelligenza alle macchine che non possono essere considerate senzienti, mancando di autocoscienza e capacità di ragionamento.

[7] S. SAPIENZA, *Decisioni algoritmiche e diritto*, cit., p.12.

[8] J. HAUGELAND, *Artificial intelligence: The very idea*, MIT press 1989; E. CHARNIAK, *Introduction to artificial intelligence*, Pearson Educaton India, 1985; R. KURWEIL, *The age of intelligence machines* vol.580 MIT press Cambridge 1990; D. POOLE, A. MACKWORTH, R. GOEBEL, *Computational Inelligence* 1990; S. RUSSELL, P. NORVIG, *Artificial intelligence: a modern approach*, Pearson Education, 2010.

[9] J. SEARLE, *Minds, brains, and programs*, in *Behavioral and brain science* 3.3, Cambridge University Press,1980, p.417.

[10] Artificial Narrow Intelligence (ANI) L'intelligenza artificiale debole che è l'unica tecnologia finora sviluppata ed applicata, si basa sul "come se" e la ritroviamo nei sistemi di simulazione di alcune funzionalità cognitive dell'essere umano, non raggiungendo, però, le capacità intellettuali tipiche dell'uomo. L'intelligenza artificiale debole, quindi, comprende programmi di problem-solving in grado di replicare alcuni ragionamenti logici umani per risolvere problemi o prendere decisioni come, ad esempio, la traduzione automatica dei testi.

[11] Artificial General Intelligence (AGI)

[12] N. BOSTROM, *Superintelligence*, Dunod 2017.

[13] S. SAPIENZA, *Decisioni algoritmiche e diritto*, cit., p.13.

(AI HLEG) della Commissione Europea[14], ad esempio, ha stabilito che i sistemi di intelligenza artificiale sono software in grado di agire e interagire con una dimensione fisica o digitale attraverso l'acquisizione e l'interpretazione di dati, di ragionare sulla base della conoscenza e di decidere per raggiungere un determinato obiettivo.

Infine, l'approccio giuridico tiene conto sia dell'aspetto funzionale, che di quello eziologico e, come previsto dall'art.3.1 del Regolamento del Parlamento Europeo e del Consiglio 2024/1689 che stabilisce regole armonizzate sull'intelligenza artificiale (*AI ACT*) per "*sistema di IA*" si intende: «*un sistema automatizzato progettato per funzionare con livelli di autonomia variabili e che può presentare adattabilità dopo la diffusione e che, per obiettivi espliciti o impliciti, deduce dall'input che riceve come generare output quali previsioni, contenuti, raccomandazioni o decisioni che possono influenzare ambienti fisici o virtuali*».

Tale definizione è stata, altresì, recepita dall'art.2 del d.d.l. AS1146 presentato dal Governo italiano al Parlamento il 20 maggio 2024, in corso di esame in Commissione[15].

Anche l'art.2 della Convenzione quadro sull'intelligenza artificiale, i diritti umani, la democrazia e lo Stato di diritto adottata dal Consiglio d'Europa[16] il 17 maggio 2024 a Strasburgo, stabilisce che «*per "sistema di intelligenza artificiale" si intende un sistema basato su una macchina che, per obiettivi espliciti o impliciti, deduce, dagli input che riceve, come generare output quali previsioni, contenuti, raccomandazioni o decisioni che possono influenzare le condizioni fisiche o ambienti virtuali. Diversi sistemi di intelligenza artificiale variano nei loro livelli di autonomia e ad attività dopo l'implementazione*».

1.1.1 Le origini dell'intelligenza artificiale

L'uomo ha sempre aspirato ad imitare e riprodurre sé stesso e la natura, seppur in forme che, solo di recente, sono quelle dell'intelligenza artificiale. Infatti, già nel mondo

[14]AI HLEG, *A definition of AI: Main Capabilities and scientific disciplines* 2019. ec.europa.eu. L'High-Level Expert Group on Artificial Intelligence (AI HLEG) è stato istituito nel giugno del 2018 ed è composto da cinquantadue esperti del mondo delle imprese, della ricerca, dell'università e della pubblica amministrazione, incaricati di proporre analisi e indicazioni per le politiche europee sull'AI.

[15] D.d.l. AS 1146 recante Disposizioni e delega al Governo in materia di intelligenza artificiale, www.senato.it

[16] Convenzione quadro sull'intelligenza artificiale, i diritti umani, la democrazia e lo Stato di diritto https://search.coe.int/

ellenistico Erone di Alessandria[17] elaborò gli automi serventi per dimostrare basilari principi scientifici di idraulica, pneumatica e meccanica.

Nel Medioevo Raimondo Lullo[18], cercò di imitare le facoltà di pensiero dell'uomo e progettò l'*Ars Magna*, una macchina composta da cerchi concentrici realizzati con dischi di metallo o gesso in cui erano inseriti numeri o simboli che rappresentavano i principi fondamentali di tutte le scienze. Attraverso la combinazione dei cerchi concentrici, Lullo riteneva si potessero riprodurre ragionamenti tipici per risolvere problemi comuni.

G. W. Leibniz alla fine del 1600 affrontò il tema della meccanizzazione del ragionamento ricercando una lingua universale e considerò il ragionamento come una sorta di algebra del pensiero. Egli riteneva che "*si può assegnare il suo numero caratteristico*" anche ai pensieri umani. In tal modo era possibile creare un alfabeto dei pensieri e "*dalla combinazione delle lettere di questo alfabeto e dall'analisi dei vocaboli formati da quelle lettere, si potevano scoprire e giudicare tutte le cose*"[19], usando un calcolo infallibile. Attraverso il *calculus ratiocinator*, dunque, sarebbe stato possibile affrontare ogni problema e risolverlo, eliminando dispute e polemiche.

Inoltre, Leibniz partendo dalla calcolatrice elaborata da Pascal per aiutare nei suoi conti il padre incaricato della ripartizione delle tasse in Normandia, nel 1671, concepì una calcolatrice meccanica destinata a eseguire somme e prodotti, che fu realizzata nel 1694.

Nel XVIII secolo si iniziò a pensare a produrre macchine che, oltre ad essere efficienti, rappresentassero effettivamente qualcosa di umano[20]. Pertanto, Jacques de Vaucanson costruì un'anitra artificiale che nuotava e ingoiava il grano. Pierre Jacquet-Droz, invece, realizzò alcuni automi, tra cui uno scrivano, un disegnatore e una musicista. I loro

[17] Erone di Alessandria era un matematico e fisico vissuto, secondo alcuni, nel III sec. d.C. che scrisse il trattato *Automata* dedicato alle macchine semoventi (da automatos, "*che agisce di propria volontà*"), noto anche con il titolo "*Sulla fabbricazione degli automi*". Esso fu tradotto nel XVI sec. Nel trattato è descritto dettagliatamente un teatrino (Automata, 4, 1-3), l'unico congegno automatico antico di cui si hanno notizie.

[18] Raimondo Lullo era un filosofo, teologo, mistico e missionario nato a Palma di Maiorca nel 1233-1235 che intendeva convertire gli ebrei e gli islamici al cristianesimo. A tale scopo elaborò una logica universale, capace di scoprire e dimostrare la verità partendo dai termini semplici combinati in modo matematico. La sua opera ebbe larga influenza sino al XVII secolo.

[19] G.W. LEIBNIZ, *Historia et commendatio linguae characticae universalis quae simul sit ars inveniendi et judicando* (1679-80 ca.) ID., *Scritti di logica*, a cura di F. BARONE, trad. it., Bologna, 1968, p.508.

[20] V.MARCHIS, *Uomini, macchine, automi*, in *Storia della civiltà europea* a cura di Umberto Eco Treccani, Roma, 2014

movimenti erano controllati attraverso dei codici memorizzati su alcuni dischi di metallo.

Nel XIX secolo il matematico e filosofo britannico Babbage, invece, elaborò due progetti: la macchina alle differenze[21] e la macchina analitica[22]. Egli, per operare con i numeri e le regole intendeva tradurli in qualcosa di meccanico, come schede e ingranaggi che, muovendosi, avrebbero prodotto una situazione tale da poter essere letta e ritradotta in numero, secondo la successione numero, stato fisico, numero.

Tuttavia, il contributo decisivo all'elaborazione di ciò che, nel 1956, per la prima volta, fu definito "*intelligenza artificiale*" è stato fornito dalla cibernetica e dall'informatica.

La cibernetica nacque negli anni '40 del 1900 per studiare i meccanismi dell'autoregolazione e del controllo presenti, sia negli organismi viventi, sia nelle macchine, che permettono di adattare il proprio comportamento alle sollecitazioni dell'ambiente[23] .

Essa si basa sull'idea che il funzionamento delle macchine costruite dall'uomo, come gli esseri viventi è regolato dalla retroazione o principio del feedback, secondo cui ad un certo stimolo – input – corrisponde una risposta – output –che altera lo stimolo, producendo un nuovo input, che causa un nuovo output, fino ad ottenere un bilanciamento generale del sistema.

Successivamente, tra il 1943 e il 1945 Warren McCulloch e Walter Pitts[24], estesero tale metodo di analisi alle funzioni cognitive del sistema nervoso centrale e alle nuove macchine calcolatrici digitali. Essi elaborarono un modello di neurone artificiale in

---

[21] V.MARCHIS, M. CORSI, *Macchine* (voce), in *Enciclopedia delle scienze sociali*, Treccani Roma, 1996. Nella macchina alle differenze si possono rappresentare fino a venti numeri di diciotto cifre ciascuno; le cifre sono rappresentate da ruote e le ruote, per le cifre di uno stesso numero, sono incolonnate in modo tale che ogni colonna possa comprendere fino a diciotto ruote.

[22] V.MARCHIS, M. CORSI, *Ivi*,. Nella macchina analitica, che non fu mai realizzata per problemi economici, i numeri risultanti come uscita della macchina alle differenze avrebbero dovuto essere utilizzati come dati in ingresso per un calcolo successivo. La parte centrale era costituita dalla "*macina*" una sorta di processore in grado di elaborare i numeri che le venivano inviati per produrre i risultati richiesti. La macchina analitica, inoltre, avrebbe dovuto possedere un magazzino (memoria) contenente sia i numeri non ancora elaborati dalla macina, sia quelli risultanti dalle operazioni della macina stessa. La macchina analitica, infine, avrebbe dovuto essere dotata anche di un sistema di controllo, in grado di eseguire automaticamente una sequenza di operazioni.

[23] A. CARACCIOLO DI FORINO, *Cibernetica* (Voce), in *Enciclopedia Italiana*, III Appendice Treccani Roma, 1961. Il termine "*cibernetica*" fu impiegato per la prima volta nel 1948 da Norbert Wiener, matematico statunitense, nel saggio *La cibernetica: controllo e comunicazione nell'animale e nella macchina* (*Cybernetics, or control and communication in the animal and the machine*) per definire un nuovo ambito di studi comune tra l'ingegneria, la biologia e le scienze umane, che si basa sulla comparazione di animale e macchina, e sul rapporto tra naturale e artificiale.

[24]J.D. COWAN, *Storia dei concetti e delle tecniche nella ricerca sulle reti neurali,* in *Frontiere della Vita,* Treccani Roma, 1999.

grado di eseguire delle funzioni logiche basilari. Come il neurone biologico, il neurone di McCulloch-Pitts riceve dei segnali in ingresso e può, eventualmente, trasmettere una risposta in uscita in base a parametri interni preimpostati. I neuroni artificiali possono essere organizzati e connessi in reti di diversa complessità formando un sistema che, secondo i ricercatori, avrebbe dovuto imparare come il cervello umano, con l'esperienza, in modo autonomo, sulla base dei tentativi e degli errori.

Nel 1949 D. O. Hebb ipotizzò che le macchine potessero essere istruite con un apprendimento simile a quello umano[25].

Sulla base degli studi di McCulloch e Pitts, Frank Rosenblatt nel 1958 progettò il Perceptron, un modello base di rete neurale[26]. Si trattava della prima macchina in grado di simulare il funzionamento dei neuroni con connessioni analoghe alle sinapsi di un neurone biologico, attraverso un software e un hardware, e di realizzare una forma elementare di apprendimento automatico. La macchina era in grado di distinguere e classificare figure a forma di lettere. Il sistema aveva un'unità di entrata (input), una di uscita (output) e si basava su una regola di apprendimento automatico fondata sulla minimizzazione dell'errore: l'algoritmo di retropropagazione dell'errore[27]. Il percettrone, dunque, a differenza del neurone di Pitts e McCullogh, aveva capacità di memorizzazione e di apprendimento.

Come si è detto, un contributo fondamentale agli studi sull'intelligenza artificiale è stato, altresì, fornito, dall'informatica, disciplina nata negli anni '50 del 1900 per studiare la rappresentazione, l'organizzazione e il trattamento automatico dell'informazione, attraverso l'elaboratore o calcolatore[28]. Nel calcolatore, infatti, è installato un programma che contiene una serie di istruzioni grazie alle quali esso può elaborare i dati inseriti dall'utente ed eseguire calcoli logico-matematici, generando informazioni organizzate.

L'informatica si basa sugli studi elaborati da John von Neumann cui si deve l'organizzazione concettuale del calcolatore moderno e di Alan Turing che studiò la

[25] J.D. COWAN, *Ivi.* Donald Hebb elaborò la "*Regola di Hebb*" secondo cui se due neuroni collegati fra loro sono attivi contemporaneamente il valore sinaptico delle loro connessioni viene aumentato.
[26] J.D. COWAN, *Ivi.*
[27] L'algoritmo confronta il valore in uscita del sistema con il valore desiderato (obiettivo). Sulla base della differenza così calcolata (errore), l'algoritmo modifica i pesi sinaptici della rete neurale.
[28] G. MARTINOTTI, *Informatica* (voce), in *Enciclopedia delle scienze sociali*, Roma, Treccani, 1994. Il termine è di origine francese e fu coniato nel 1962 dall'ingegnere Philippe Dreyfus, per composizione dei vocaboli "*information*" e "*automatique*". Esso fu impiegato per indicare il trattamento razionale, con macchine automatiche, dell'informazione, intesa come supporto delle conoscenze umane.

crittografia ed elaborò la macchina di Turing[29]. Tale macchina è composta da un nastro di lunghezza infinita, suddiviso in celle lette da una testina, che può spostarsi di una cella avanti o indietro, e da un organo di controllo capace di leggere il simbolo. A un dato istante, l'azione che la macchina intraprende è determinata dal simbolo letto e dalla configurazione in cui la macchina si trova in quel momento. Dopo aver letto il simbolo stampato sulla cella, la testina può compiere due operazioni alternative: lasciare il simbolo così com'è, oppure cancellarlo e stamparne un altro.

Alan Turing nel 1950 pubblicò un articolo "*Computing machinery and intelligence*", in cui espose il test di Turing[30], necessario per stabilire se una macchina fosse in grado di avere un comportamento intelligente.

Gli studi di cibernetica e di informatica costituirono la base per il seminario del Dartmouth College che si tenne nell'estate del 1956 al quale parteciparono ricercatori interessati alla teoria degli automi, alle reti neurali e allo studio dell'intelligenza[31].

Dieci studiosi per due mesi si sarebbero confrontati sull'intelligenza artificiale partendo dalla congettura per cui, in linea di principio, ogni aspetto dell'apprendimento o una qualsiasi altra caratteristica dell'intelligenza possano essere descritte così precisamente, da poter costruire una macchina che le simuli [32].

Lo scopo del seminario era «*scoprire come far sì che le macchine utilizzino il linguaggio, formino astrazioni e concetti, risolvano tipi di problemi ora riservati agli*

---

[29] M.CAPPELLI, *Macchina di Turing* (voce), in *Enciclopedia della Scienza e della Tecnica*, Treccani Roma,2008.

[30] *Test di Turing* (voce), in *Enciclopedia della matematica*, Treccani Roma 2013. Il test coinvolgeva tre partecipanti: un uomo A, una donna B e una terza persona C che è separata dagli altri due e funge da arbitro. C, senza vedere A e B deve porre loro delle domande e, in base alle risposte dattiloscritte deve stabilire qual è l'uomo e qual è la donna. Allo stesso tempo A deve fornire risposte ingannevoli e indurre C ad un'identificazione errata, mentre B deve aiutarlo. Dopo questa prima fase, A è sostituito da una macchina. In tal caso, se C non si accorgesse di nulla, la macchina dovrebbe essere considerata intelligente. Essa, infatti, sarebbe indistinguibile da un essere umano e, quindi intelligente.

[31] F. AMIGONI, V. SCHIAFFONATI, M. SOMALVICO, *Intelligenza artificiale* (voce), in *Enciclopedia della Scienza e della Tecnica*,2008, Treccani Roma,

[32] I presupposti del seminario sono delineati nel documento "*Una proposta di progetto per una ricerca estiva a Dartmouth sull'Intelligenza artificiale"*, in *AI Magazine*, 27.4.2006, p.12 «*We propose that a 2 month, 10 man study of artificial intelligence be carried out during the summer of 1956 at Dartmouth College in Hanover, New Hampshire. The study is to proceed on the basis of the conjecture that every aspect of learning or any other feature of intelligence can in principle be so precisely described that a machine can be made to simulate it*.» Il documento fu redatto da John McCarthy e Marvin Lee Minsky, matematici statunitensi, considerati tra i padri dell'informatica e dell'intelligenza artificiale, da Claude Elwood Shannon, un matematico e ingegnere statunitense, considerato il padre della teoria dell'informazione e da Nathaniel Rochester, un ingegnere elettronico statunitense che ha sviluppato il linguaggio di programmazione a basso livello Assembly. Al seminario parteciparono altri sei studiosi: Ray Solomonoff, Oliver Selfridge, Trenchard More, Arthur Samuel, Allen Newell e Herbert Simon.

esseri umani e migliorino se stesse […]. Ai fini attuali, il problema dell'intelligenza artificiale è quello di far sì che una macchina si comporti in modi che verrebbero definiti intelligenti se un essere umano si comportasse in questo modo»[33].

Durante il seminario Allan Newell, Herbert Alexander Simon e Cliff Shaw presentarono il Logic Theorist, il primo programma di ragionamento automatico in grado di dimostrare trentotto dei cinquantadue teoremi enunciati nel trattato di Whitehead e Russell, *Principia Mathematica*. Il programma si basava sulla logica simbolica, la branca della matematica che si occupa di convertire in simboli, concetti ed affermazioni, per trasformarli e svolgere induzioni e deduzioni. Con il Logic Theorist, quindi, si cercò di imitare le capacità di "*problem solving*" degli esseri umani[34].

Gli anni immediatamente successivi al seminario furono molto entusiasmanti per la ricerca che, peraltro, ottenne consistenti finanziamenti. In questi anni, ad esempio, si lavorò all'elaborazione del linguaggio naturale e fu sviluppato ELIZA, il primo chatbot della storia. Il programma era in grado di conversare impiegando schemi di risposta che ripetevano le parole dell'interlocutore.

Alcuni ricercatori come Herbert Simon e Marvin Minsky si spinsero persino ad affermare che "*Le macchine saranno capaci, nell'arco di trent'anni, di fare qualsiasi lavoro un uomo può fare*" (H. Simon, 1956), "*Nel giro di una generazione ... il problema di creare un'intelligenza artificiale sarà sostanzialmente risolto*" (M. Minsky, 1967), "*Tra tre-otto anni avremo una macchina con l'intelligenza generale di un essere umano medio*" (M. Minsky, 1970)[35].

Nello stesso periodo fu messa in discussione la validità dei modelli a reti neurali. Infatti, Marvin Minsky e Seymour Papert nel 1969 elaborarono una critica al Perceptron di Frank Rosenblatt affermando che la macchina fosse incapace di riconoscere stimoli visivi anche molto semplici[36].

---

[33] J.MC CARTHY et.al., *Una proposta di progetto per una ricerca estiva a Dartmouth sull'Intelligenza artificiale*, cit., «*to find how to make machines use language, form abstractions and concepts, solve kinds of problems now reserved for humans, and improve themselves […]. For the present purpose the artificial intelligence problem is taken to be that of making a machine behave in ways that would be called intelligent if a human were so behaving*»

[34] F. AMIGONI, V. SCHIAFFONATI, M. SOMALVICO, *Intelligenza artificiale* (voce), cit.

[35] L.PORTINALE, *Intelligenza Artificiale: storia, progressi e sviluppi tra speranze e timori*, in *Medialaws*, 2021, 3, p.18.

[36] M. MINSKY e S. PAPERT, *Perceptrons: an introduction to computational geometry*. La pubblicazione del libro provocò una notevole riduzione dei finanziamenti per la ricerca nel campo delle reti neurali artificiali. Come si vedrà in seguito, solo nella prima metà degli anni Ottanta, la ricerca riprese vigore.

Pertanto, la ricerca abbandonò l'approccio connessionista, basato su modelli ispirati alle connessioni tra neuroni nel cervello umano, a favore di quello fondato sulla logica simbolica, e si sviluppò in due direzioni distinte. Il gruppo guidato da Allan Newell, Cliff Shaw e Herbert Alexander Simon continuò a studiare la simulazione dei processi cognitivi umani attraverso l'elaboratore, creando il GPS (General problem solver), un programma che intendeva estendere l'ambito delle applicazioni del Logic Theorist.

Vi erano, poi, altri ricercatori che intendevano potenziare le prestazioni dei programmi, anche senza adottare necessariamente strumenti che imitassero i procedimenti seguiti dall'uomo. Essi elaborarono i primi programmi per il gioco della dama e degli scacchi e per la risoluzione di problemi matematici[37].

Tuttavia, i ricercatori furono costretti a riconoscere che comprendere e replicare l'intelligenza umana è estremamente difficile. Infatti, man mano che si affrontavano problemi più complessi e ampi, emergeva l'inadeguatezza degli strumenti fino ad allora utilizzati, di fronte alla necessità di elevatissime capacità di calcolo e di memoria. L'intelligenza artificiale, di fatto, non aveva alcuna utilità in ambito scientifico o industriale e furono revocati gli ingenti investimenti pubblici e privati per la ricerca nel settore.

1.1.2 Gli inverni dell'intelligenza artificiale e i sistemi esperti

Il cosiddetto "primo inverno"[38] dell'intelligenza artificiale ebbe inizio negli anni '70 del 1900 e durò circa un decennio, finché furono creati i "sistemi esperti"[39]. Infatti, i ricercatori, accantonata l'idea di costruire un sistema di intelligenza generale e assoluta, capace di risolvere qualunque problema, portarono avanti l'approccio basato sulla conoscenza. Pertanto, cercarono di realizzare sistemi programmati non per ragionare, ma per conoscere. Si tratta di sistemi che, in un settore specifico come, ad esempio, quello medico, mettevano assieme conoscenze, non solo teoriche, ma anche pratiche,

[37] F. AMIGONI, V. SCHIAFFONATI, M. SOMALVICO, *Intelligenza artificiale* (voce), cit.

[38] M. LIM, *History of AI winters*, in *Actuaries Digital*, 2018, https://www.actuaries.digital/; S. SCHUCHMANN, *History of the first AI winter, in Towards data science*, 2019, https://towardsdatascience.com/.

[39] F. AMIGONI, V. SCHIAFFONATI, *Sistemi esperti* (voce), in *Enciclopedia della Scienza e della Tecnica*, Treccani Roma 2008.

similmente ad un esperto umano. Tali conoscenze erano elaborate da un motore di inferenza che implementava opportuni algoritmi di ragionamento[40].

Attraverso i sistemi esperti, dunque, era possibile affrontare singole problematiche usando tutto il patrimonio di conoscenze teoriche e pratiche sull'argomento. Sistemi come DENDRAL[41] elaborato dal Massachusetts Institute of Technology (MIT) e MYCIN[42] elaborato dalla Stanford University, ad esempio, si basavano sul trasferimento diretto alla macchina, di tutte le conoscenze teoriche e pratiche dell'uomo, necessarie, rispettivamente, per analizzare la struttura molecolare e per effettuare la diagnosi di infezioni sanguigne. PROSPECTOR, invece, era un sistema esperto in prospezioni geologiche sviluppato dall'SRI International che riuscì a localizzare un giacimento di molibdeno nel nord ovest degli Stati Uniti.

In questo periodo fu elaborato anche R1, il primo sistema esperto commerciale di successo che supportava le configurazioni di ordini per nuovi sistemi di elaboratori, nell'azienda informatica statunitense Digital Equipment Corporation.

I benefici economici connessi all'utilizzo dei sistemi esperti furono notevoli e numerose grandi imprese investirono nella loro realizzazione sviluppando un'industria intorno a tale forma di intelligenza artificiale[43]. Inoltre, ripresero anche i programmi di finanziamento pubblico della ricerca. Nel 1981, infatti, il Giappone avviò un programma di investimenti decennale, invitando, come partner, il Regno Unito che, tuttavia, date le enormi potenzialità del settore, decise di realizzare un autonomo piano nazionale di sviluppo tecnologico.

Anche gli Stati Uniti sostennero la creazione di un consorzio con le più importanti imprese del settore e i centri di ricerca pubblici e privati più avanzati.

Tuttavia, ben presto, anche i sistemi esperti mostrarono i propri limiti. Infatti, il problema principale di tali sistemi era costituito dall'acquisizione della conoscenza essenziale per farli funzionare ed aggiornarli. In mancanza di metodi e strumenti per

[40]Secondo la definizione di Peter Jackson, *Introduction to Expert Systems*, 3rd ed., International Computer Science Series 1999 i sistemi esperti sono «*sistemi computerizzati in grado di emulare l'attività di decisione di un essere umano esperto in un determinato settore*».

[41] Ed Feigenbaum (studente di Herbert Simon), Bruce Buchanam e Joshua Lederberg crearono il primo sistema esperto, DENDRAL, programmato per inferire la struttura delle molecole organiche in base alle loro formule chimiche.

[42] MYCIN, derivato da DENDRAL, è un sistema che usa la conoscenza medica specifica per diagnosticare e prescrivere trattamenti per le infezioni batteriche del sangue a partire da una conoscenza incompleta e incerta.

[43] Le grandi imprese americane utilizzavano sistemi esperti che generarono un volume d'affari passato, in meno di dieci anni, da pochi milioni a miliardi di dollari.

l'apprendimento automatico le conoscenze dovevano essere acquisite e aggiornate manualmente ed era necessaria la collaborazione degli esperti di ciascuno specifico settore di interesse e di scienziati specializzati nei metodi di intelligenza artificiale[44].
Tra la fine degli anni '80 e l'inizio degli anni '90 del 1900, quindi, si ebbe il secondo inverno dell'intelligenza artificiale, i fondi per la ricerca furono ridotti e migliaia di imprese cessarono o convertirono la propria attività[45].

1.1.3 L'apprendimento automatico

Negli anni '80 del 1900 diversi gruppi di ricerca diedero nuovo impulso all'approccio connessionista che si rivelò molto promettente[46]. I ricercatori ripresero lo studio sulle reti neurali e, applicando i più recenti risultati della matematica e della fisica, elaborarono un algoritmo di apprendimento basato sulla retropropagazione dell'errore. In particolare, John Hopfield creò un modello di rete neurale che produceva memoria c.d. associativa e che, impiegando le conoscenze accumulate con l'esperienza, operava in presenza di informazioni scarse o deteriorate[47].
A partire dagli anni '90 del 1900 i ricercatori riuscirono ad ottenere anche un incremento della capacità computazionale dei computer. Inoltre, lo studio dell'intelligenza artificiale divenne più scientifico e si integrò con altre discipline, per cercare di migliorare teorie già formulate e tecnologie già presenti, anziché crearne di nuove. Pertanto, assunsero particolare rilevanza la teoria della probabilità, gli studi sul linguaggio e i risultati ottenuti dalle scienze matematiche, statistiche ed economiche sui procedimenti di decisione e di scelta razionale.
In questo periodo i ricercatori si dedicarono alla creazione di agenti software intelligenti e di agenti intelligenti "incorporati" in un sistema fisico, cioè dei robot con capacità cognitive, che si interfacciavano con il mondo esterno e potevano cambiarlo. Essi

[44] T.M. LAZZARI, F.L. RICCI, *I sistemi esperti. Ricerca scientifica ed applicazioni*, Nuova Italia Scientifica, Roma 1985.
[45] M. LIM, *History of AI winters* cit.
[46] Si vedano ad esempio i lavori, condotti da D.E. RUMELHART, G.E. HINTON, R.J. WILLIAMS, *Learning representations by back-propagating errors*, in *Nature*, 323, 6088, 1986, p. 533–536; D.B. PARKER, *Learning Logic*, MIT Center for Computational Research in Economics and Management Science – Technical Report, 1985; Y. LECUN, *Une procedure d'apprentissage pour réseau a seuil asymmetrique* (*a Learning Scheme for Asymmetric Threshold Networks*), in *Proceedings of Cognitiva* 85, Paris,1985, p. 599-604.
[47] F. AMIGONI, V. SCHIAFFONATI, M. SOMALVICO, *Intelligenza artificiale* (voce), cit.

condivisero le strategie, i risultati e i dati necessari per lavorare e per valutare i progressi ottenuti[48].

Negli anni 2000, la nascita del World Wide Web e l'introduzione sul mercato di chip di elaborazione dati molto veloci consentirono l'accesso a notevoli quantità di informazioni e conoscenze, incentivando lo sviluppo di nuovi algoritmi e applicazioni per l'apprendimento automatico. Con i numerosi dati disponibili era possibile addestrare un sistema informatico ad eseguire attività, senza necessità di fornirgli istruzioni esplicite e regole, ottenendo risultati anche per scenari sconosciuti o nuovi. Pertanto, furono elaborate metodologie per il riconoscimento di oggetti, persone e luoghi, per comprendere e tradurre il linguaggio umano, per analizzare i comportamenti o le preferenze di utenti, per prevedere e prevenire richieste.

I ricercatori evidenziarono che, aumentando la grandezza dei dataset su cui addestrare gli algoritmi era possibile raggiungere risultati assimilabili alle performance di un individuo. La qualità del sistema di elaborazione dei dati, dunque, era considerata meno importante della quantità dei dati disponibili.

L'apprendimento automatico o *machine learning*[49], infatti, si basa su algoritmi addestrati ad inferire dei modelli, individuati da un set di dati, per determinare le azioni necessarie a raggiungere un obiettivo. Esso non richiede una preventiva programmazione del sistema che apprende dall'esperienza.

L'apprendimento è supervisionato quando agli algoritmi sono forniti i dati in entrata e i risultati da raggiungere. Il compito dell'intelligenza artificiale consiste nel collegare i dati e i risultati e individuare il nesso logico, da applicare ad altri casi analoghi[50].

L'apprendimento è non supervisionato, invece, quando gli algoritmi sono addestrati solo con dati in ingresso e devono individuare la struttura logica che consente di raggiungere un risultato.

L'apprendimento semi-supervisionato combina l'apprendimento supervisionato e non supervisionato.

L'apprendimento per rinforzo, infine, è il sistema di apprendimento più complesso e richiede che la macchina sia dotata di strumenti e sistemi in grado di migliorare il

[48] L.PORTINALE, *Intelligenza Artificiale: storia, progressi e sviluppi tra speranze e timori*, cit., p.20.
[49] A. L. SAMUEL, *Some studies in machine learning using the game of checkers*. II *Recent progress in IBM Journal of research and development*, II.6, 1967, p.601. La creazione dell'espressione machine learning è attribuita a tale autore.
[50] S. SAPIENZA, *Decisioni algoritmiche e diritto*, cit., p.17.

proprio apprendimento e di comprendere le caratteristiche dell'ambiente in cui opera. Gli algoritmi, infatti, scelgono una determinata azione sulla base della percezione dell'ambiente con cui interagiscono. Essi sono penalizzati o ricompensati a seconda della scelta effettuata. L'algoritmo tende a raggiungere la maggiore quantità di punti di ricompensa possibile (*fitness*) definito dagli sviluppatori e, eventualmente, a realizzare un obiettivo finale.

La diffusione dell'utilizzo di internet ha determinato il vertiginoso aumento dei dati disponibili anche per l'apprendimento. Nel 2001 Doug Laney[51] per la prima volta usò l'espressione "*Big Data*" per definire un contesto di elaborazione dei dati dove si hanno grandi volumi, si elabora a grandi velocità, si ha una grande varietà di dati.

La locuzione fu nuovamente impiegata nel 2011 nel rapporto intitolato "*Big Data: The next frontier for innovation, competition and productivity*"[52] per indicare la profonda trasformazione sociale, economica ed umana che aveva reso disponibile una grande quantità di dati da acquisire, processare, analizzare e valorizzare.

Con l'espressione "*Big data*", dunque, si indica l'accumulo sistematico di enormi quantità di dati, all'interno di computer centralizzati con grandi capacità di memoria e di calcolo. Tali dati derivano dalla digitalizzazione di fonti tradizionali come biblioteche, archivi di Stato, oppure sono prodotti dalla navigazione in rete o sono le foto e i video condivisi in tempo reale sui social o sulle app di messaggistica istantanea. Inoltre, vi sono i dati raccolti da sensori presenti nell'ambiente e connessi ad internet, capaci di fornire a chi li raccoglie ogni genere di informazione sulla vita di ciascuno.

L'esplosione dei *Big data* ha dato ulteriore impulso al *machine learning*, nell'ambito del quale si è sviluppata una nuova forma di addestramento delle reti neurali modellata sul cervello umano, il *deep learning*[53]

Tale modalità di apprendimento, infatti, utilizza le reti neurali artificiali per elaborare le informazioni su più livelli. Il primo strato di neuroni della rete neurale riceve i dati in ingresso dall'esterno, mentre gli strati successivi utilizzano i dati risultanti dallo strato precedente. Più la rete è profonda, più l'informazione è elaborata in modo completo e più l'output del sistema è accurato. Le applicazioni del *deep learning* sono numerose e

[51] D.LANEY, *3D Data Management: Controlling Data Volume, Velocity, and Variety*, Meta Group 2001.
[52] J. MANYIKA, M.CHUI, B. BROWN, J. BUGHIN, R. DOBBS, C. ROXBURGH, A. BYERS, *Big Data: The next frontier for innovation, competition and productivity*, McKinsey Global Institute, 2011.
[53] L.PORTINALE, *Intelligenza Artificiale: storia, progressi e sviluppi tra speranze e timori*, cit., p.22.

hanno consentito la creazione di sistemi di riconoscimento vocale, di traduzione automatica, di diagnostica per immagini o di decisione automatica.

Il Large Language Model (LLM) come Chat GPT[54], ad esempio, è basato sul *deep learning* e consente di comprendere e generare testi in modo autonomo. Il LLM è una rete neurale profonda addestrata su un notevole numero di libri, articoli, pagine web e altre fonti tratte dal Web. Il modello impara progressivamente a riconoscere le strutture linguistiche dei testi in modo da poter predire la parola che seguirà un certo input che è detto "*prompt*".

Il modello in esame, dunque, si differenzia dai motori di ricerca in quanto non si limita ad ottimizzare la ricerca in un database, bensì genera risposte alle domande, sulla base di un proprio set di addestramento, utilizzando correlazioni meramente statistiche. Tale modello, infatti, associa le parole convertite in matrici, non per il loro valore semantico, bensì per la frequenza statistica con cui esse sono combinate nei data set di addestramento, adottando un approccio probabilistico e non deterministico. Pertanto, esso può generare dei risultati che, pur essendo statisticamente appropriati, non sono eziologicamente e ontologicamente corretti.

La ricerca sull' intelligenza artificiale, dunque, segue due approcci distinti.

Secondo l'approccio simbolico o Model Based AI i sistemi di intelligenza artificiale sono modelli rigidi, programmati a priori con tutte le possibili combinazioni di fattori utili per la soluzione del problema, sulla base di regole logico deduttive. Il loro vantaggio è dato dalla affidabilità dei risultati che sono spiegabili e controllabili. Tuttavia, il loro limite è costituito dall'impossibilità di rappresentare in essi tutta la conoscenza.

Secondo l'approccio sub simbolico o connessionista o Machine Learning AI, basato sull'apprendimento, invece, i sistemi di intelligenza artificiale non ricevono alcuna programmazione o regola logica ma funzionano attraverso l'addestramento con una grande quantità di dati. Tali dati sono trattati secondo la computazione probabilistica, un paradigma flessibile basato sulla inferenza-abduzione che genera come risultato la risposta più credibile. Tale approccio consente di ottenere i modelli in modo semplice e naturale. Tuttavia, i sistemi basati sull'apprendimento producono l'effetto cosiddetto della Scatola Nera (*Black Box*), cioè non sono in grado di dare una spiegazione del

[54] S. SAPIENZA, *Decisioni algoritmiche e diritto*, cit., p.22. ChatGPT è un software sviluppato da OpenAI.

risultato ottenuto. Inoltre, essi sono influenzati dai pregiudizi (*Bias*) causati dai dati di addestramento che possono essere parziali e indurre a conclusioni non neutrali e oggettive. I modelli in esame non hanno la percezione del rapporto tra causa ed effetto e possono essere facilmente tratti in errore. Infine, essi presentano dei rischi anche per la tutela della riservatezza dei dati utilizzati per l'apprendimento[55].

Attualmente i sistemi di intelligenza artificiale sono in grado di processare cioè di interpretare una mole enorme di dati che sono inseriti dal programmatore oppure che la macchina può autonomamente apprendere, possono cogliere, percepire, alcune caratteristiche dell'ambiente in cui sono inseriti, riescono a dare risultati in forma variabilmente autonoma e imprevedibile e, quindi, possono decidere e, infine, sono in grado di eseguire compiti relativamente complessi, anche imparando, cioè modificando il proprio funzionamento per migliorare le proprie prestazioni.

1.2 L'intelligenza artificiale e il diritto

La riflessione in merito al rapporto tra diritto ed informatica[56] ha avuto inizio sin dagli anni '50 del 1900 nell'ambito della trasformazione delle metodologie di ricerca scientifica sui rapporti fra l'uomo e gli oggetti e sui rapporti fra gli uomini stessi.

Lee Loevinger[57] propose di utilizzare i metodi della scienza nel campo del diritto, applicando la nuova tecnica dell'automazione e dell'elaborazione elettronica dei dati. Egli utilizzò per la prima volta il termine "*giurimetria*" per indicare l'uso degli elaboratori informatici per "misurare" le decisioni giudiziarie e verificarne la prevedibilità, attraverso i precedenti giurisprudenziali. Partendo dal presupposto che il diritto è uno strumento misurabile, calcolabile e, pertanto, suscettibile di essere tradotto nel linguaggio informatico del calcolatore, la giurimetria intendeva studiare e schematizzare la logica ricorrente nel diritto e nella decisione giudiziaria, in modo da trarne modelli matematici e formulare, in base ad essi, previsioni attendibili che garantissero la certezza del diritto.

---

[55] S. SAPIENZA, *Decisioni algoritmiche e diritto*, cit., p.29; P. TRAVERSO, *Breve introduzione tecnica all'Intelligenza Artificiale*, DPCE online, 2022, 1, p.164.

[56] A. MANGIAMELI, *Informatica giuridica: appunti e materiali ad uso di lezioni*, Giappichelli, Torino 2015; R. BORRUSO, *L'informatica del diritto*, Giuffrè, Milano 2007; G. SARTOR, *L'informatica giuridica e le tecnologie dell'informazione. Corso d'informatica giuridica*, Giappichelli, Torino 2016; G. ZICCARDI, *Informatica giuridica: manuale breve*, Giuffrè, Milano 2011.

[57] L.LOEVINGER, *Jurimetrics. The next step forward*, in *Minnesota Law Review*, 1949, p.455.

Il diritto, dunque, era considerato come la matematica e la decisione del giudice era il frutto di una computazione[58].

La diffusione degli studi di Norbert Wiener sulla cibernetica indusse il filosofo Mario Giuseppe Losano a proporre di sostituire il termine "*giurimetria*" con il termine "*giuscibernetica*", per indicare un nuovo modo di affrontare i problemi del diritto[59].

Il giurista Vittorio Frosini[60], invece, suggerì l'adozione del termine "*giuritecnica*" per indicare la «*la tecnologia giuridica, e cioè la -produzione in atto delle -metodologie operative nel campo del diritto risultanti dall'applicazione di procedimenti e di strumenti tecnologici*».

Tali termini furono ben presto sostituiti dall'espressione generica *"informatica giuridica"*. La prima raccomandazione "*Informatica e diritto*" adottata dal Consiglio d'Europa il 30 aprile 1980[61] precisa che l'informatica giuridica è «*il ricorso alle tecnologie informatiche e di comunicazione per migliorare l'accesso dei cittadini alla giustizia e l'efficacia dell'azione giudiziaria intesa come attività di ogni genere per risolvere una controversia o sanzionare penalmente un comportamento*».

Come rilevato da autorevole dottrina[62], dunque, l'informatica giuridica «*si concentra sugli strumenti informatici a disposizione del giurista, sulla redazione di ipertesti giuridici e, più in generale, sulle applicazioni utili per la documentazione e la consultazione nell'ambito della giurisprudenza*» e va distinta dal diritto dell'informatica che, invece, studia le vicende giuridiche nel contesto telematico: la libertà di comunicazione, la tutela dei dati personali, la rilevanza giuridica del documento informatico e delle firme elettroniche, la formazione e conclusione dei contratti del commercio elettronico, la proprietà intellettuale nella distribuzione elettronica, la responsabilità civile degli operatori in rete.

L'introduzione di tecnologie digitali nel mondo del diritto e, in particolare, nel settore della giustizia è ormai consolidata. L'emergenza causata dalla pandemia da Covid 19 ha contribuito ad accelerare il processo di digitalizzazione dell'amministrazione della

[58] L.VIOLA, *Giurimetria* (voce), in *Enciclopedia on line*, Treccani Roma, 2020.
[59] M. G. LOSANO, *Giuscibernetica. Macchine e modelli cibernetici nel diritto*, Einaudi, Torino, 1969.
[60] V.FROSINI, *La giuritecnica: problemi e proposte*, in *Informatica e diritto*, I, 1975, p.6.
[61] Raccomandazione citata da R. BORRUSO, C. TIBERI, *L'informatica per il giurista. Dal bit ad internet*, II ed., Giuffrè Milano, 2001, p.49.
[62] M.G. LOSANO, *Informatica Giuridica* (voce), in *Dig. civ.*, IX, UTET, Torino, 1993, p.417.

giustizia e sono stati implementati il Processo Amministrativo Telematico, il Processo Tributario Telematico, il Processo Civile Telematico e il Processo Penale Telematico[63].
In ambito giudiziario la digitalizzazione è stata realizzata, in primo luogo, con la dematerializzazione che consiste nella trasformazione dei documenti redatti in cartaceo, in una forma che consenta di elaborarli con i computer. Inoltre, sono stati creati atti e provvedimenti nativamente digitali che possono formare oggetto di trattamento da parte di specifici software. In tal modo il testo degli atti e dei provvedimenti può essere trattato e possono essere realizzati banche-dati e modelli. La dematerializzazione permette, altresì, di celebrare da remoto le udienze.
Tuttavia, vi è anche una forma di digitalizzazione della giustizia che consiste nella conversione degli atti, dei provvedimenti, delle persone, degli spazi e dei tempi in Big data che il computer può trattare in maniera aggregata attraverso specifici software, in modo automatico.
Tali dati sono elaborati con tecniche di *machine learning* che permettono di estrarre informazioni utili per l'amministrazione della giustizia e per effettuare previsioni e calcoli di convenienza circa gli esiti di una controversia e indirizzare le decisioni del giudice[64].
Anche in ambito giuridico, dunque, si possono individuare due modelli di intelligenza artificiale. Il modello forte emula le capacità cognitive dell'uomo e postula l'automazione del processo decisionale in luogo degli attori tradizionali della giurisdizione; il modello debole, invece, è collaborativo e consente all'uomo di mantenere il controllo sull'algoritmo[65].
Attualmente, in ambito giudiziario e legale trovano applicazione soprattutto i sistemi di intelligenza artificiale debole, attraverso i programmi per la redazione o il controllo di contratti e documenti, per le valutazioni tecniche e le banche dati di giurisprudenza. Tali sistemi sono basati su algoritmi predefiniti e sono incapaci di apprendimento autonomo di nuovi dati. Essi stanno rivoluzionando anche la professione legale, consentendo un accesso agevolato a enormi quantità di dati associabili in tempo reale.

---

[63] A. GARAPON, J. LASSEGUE, *La giustizia digitale. Determinismo tecnologico e libertà*, Il Mulino, Bologna 2021.
[64] F. COSTANTINI, *Giustizia elettronica e digitalizzazione giudiziale: contesto europeo ed esperienza italiana*, in T.Casadei, S. Pietropaoli, *Diritto e tecnologie informatiche. Questioni di informatica giuridica, prospettive istituzionali e sfide sociali*, Wolters Kluwer, Milano, 2021.p. 106.
[65] G. CANZIO, *Intelligenza artificiale e processo penale*, in G. Canzio, L. Lupária (a cura di), *Prova scientifica e processo penale*, CEDAM, Padova, 2022, p.903.

Tuttavia, sono stati creati anche sistemi di polizia predittiva, che permettono di prevedere il momento e il luogo in cui determinati reati potrebbero essere commessi oppure persino chi potrebbe commetterli. Essi si basano sull'elaborazione di grandi quantità di dati riguardanti i reati già commessi, le attività delle persone sospettate, l'origine etnica e il livello di scolarizzazione dei soggetti, i luoghi a maggiore tasso di criminalità e persino le condizioni atmosferiche maggiormente correlate alla commissione di reati[66].

L'intelligenza artificiale è impiegata anche per prevenire o comporre liti e risolvere controversie, risparmiando sui tempi e sui costi dei metodi tradizionali.

Alcuni sistemi di intelligenza artificiale, come ad esempio, quelli di riconoscimento facciale automatico, trovano impiego nella ricerca della prova nella fase delle indagini preliminari e sono usati per acquisire «*tracce, cose [e] notizie idonee ad assumere rilevanza probatoria*»[67]. Tali sistemi, peraltro, possono essere d'ausilio anche nella valutazione dell'affidabilità degli elementi probatori nel processo.

L'intelligenza artificiale può essere utilizzata per la giustizia predittiva, cioè per formalizzare decisioni, parziali o totali, attraverso l'uso di un computer che apprende dall'esperienza[68]. Altre applicazioni di intelligenza artificiale permettono di adottare decisioni automatizzate in determinate fasi processuali, come la fissazione della cauzione o la determinazione delle pene.

Attraverso l'intelligenza artificiale, dunque, è possibile rendere conoscibili e trasparenti le norme e la giurisprudenza, eliminare le asimmetrie informative\relazionali, ridurre il peso dell'intermediazione legale e risolvere il problema del contrasto giurisprudenziale, garantendo la certezza del diritto e la parità di trattamento. Inoltre, è possibile ridurre l'arretrato giudiziario, realizzare risparmi economici ed economie di scala, ridurre la durata dei processi e migliorare la programmazione giudiziaria e la valutazione dei risultati raggiunti.

Tuttavia, a fronte delle opportunità connesse all'utilizzo dell'intelligenza artificiale, occorre considerare le minacce che sono insite in tale tecnologia.

---

[66] F. BASILE, *Intelligenza artificiale e diritto penale: quattro possibili percorsi di indagine*, in *Diritto Penale e Uomo*, 10, 1, 2019 p.10.

[67] L. ROMANO, *Intelligenza artificiale come prova scientifica nel processo penale*, in G. Canzio, L. Lupária (a cura di), *Prova scientifica e processo penale*, cit., p.925.

[68] M. BARBERIS, *Giustizia predittiva: ausiliare e sostitutiva. Un approccio evolutivo*, in *Milan Law Review*,2,1,2022 p.17; C. BARBARO, *Uso dell'Intelligenza Artificiale nei sistemi giudiziari: verso la definizione di principi etici condivisi a livello europeo?* in *Questione Giustizia*, 4, 2018.

Come si è detto in precedenza, infatti, i sistemi di intelligenza artificiale basati sull'apprendimento non permettono di avere una spiegazione dei risultati ottenuti partendo dagli input ricevuti. Inoltre, può accadere che i dati utilizzati per l'apprendimento siano imprecisi, viziati, poco accurati e carenti e ciò può portare a risultati di scarsa qualità, basati su pregiudizi discriminatori e lesivi dei diritti fondamentali degli individui.

Occorre, altresì, considerare che i sistemi di intelligenza artificiale in esame sono sviluppati da soggetti privati che attraverso essi detengono il controllo sui dati degli individui, acquisendo il potere di sorvegliare e di manipolare le decisioni legate alla loro sfera privata e alla dimensione pubblica. L'intelligenza artificiale, dunque, può diventare funzionale a favorire il perseguimento e il raggiungimento di interessi e di obiettivi di soggetti terzi, senza che le persone interessate siano consapevoli dell'influenza esercitata sulla loro libertà decisionale e senza che esse abbiano strumenti per tutelare i propri diritti.

Occorre, pertanto, rilevare che l'intelligenza artificiale, oltre ad essere uno strumento utile per l'amministrazione della giustizia, è essa stessa oggetto del giudizio[69]. Infatti, è necessario adattare le categorie giuridiche tradizionali ad una nuova realtà, in cui le decisioni sono assunte non da persone umane, ma da algoritmi. Si pensi, ad esempio, al problema dell'imputazione della responsabilità derivante da lesioni a beni giuridicamente protetti causate dalle auto a guida autonoma oppure dai droni o dai sistemi automatizzati di diagnosi e di cura delle malattie. In tali casi, infatti, categorie giuridiche tradizionali come la responsabilità civile, la quantificazione dei danni, la distribuzione dei diritti e degli obblighi legati a determinati comportamenti e scelte, non sempre risultano applicabili.

1.3 L'intelligenza artificiale nel diritto positivo: il Regolamento del Parlamento e del Consiglio 2024/1689 del 21 maggio 2024 (*AI Act*) e le Direttive sulla Responsabilità civile.

L'intelligenza artificiale, come si è detto, ha una molteplicità di applicazioni attuali e potenziali in grado di trasformare tutti gli ambiti della vita economico-sociale della

[69] G. DI ROSA, *Quali regole per i sistemi automatizzati "intelligenti"?* in *Rivista di Diritto Civile*, 2021, p.823; A. D'ALOIA (a cura di), *Intelligenza artificiale e diritto: come regolare un mondo nuovo*, Milano, 2020.

collettività e dei singoli individui, con un impatto notevole sui valori e sui diritti fondamentali.

Appare, dunque, necessario l'intervento del diritto per la soluzione delle numerose questioni aperte dall'avvento di tale tecnologia, che rappresentano alcune tra le principali sfide giuridiche del presente e del futuro.

Le potenze economiche mondiali[70], in primo luogo, hanno individuato le linee guida dello sviluppo dell'intelligenza artificiale nei prossimi decenni, attraverso strumenti programmatici finalizzati a tutelare l'interesse nazionale o sovranazionale. Non sono mancate, poi, dichiarazioni non vincolanti e raccomandazioni che indicano i principi per orientare lo sviluppo tecnologico verso un'intelligenza artificiale etica, affidabile e antropocentrica, elaborati da organismi internazionali, gruppi di esperti, centri di ricerca e organizzazioni non governative.

L'Unione Europea (UE), invece, ha introdotto la prima disciplina giuridicamente vincolante attraverso il Regolamento del Parlamento e del Consiglio 2024/1689 entrato in vigore a partire dal 2 agosto 2024.

Nel 2018 l'UE rese pubblica la sua strategia per l'intelligenza artificiale, con la Comunicazione della Commissione al Parlamento Europeo e al Consiglio *L'intelligenza artificiale per l'Europa*. Il piano si basava su tre pilastri: consolidare l'UE come potenza tecnologica e incoraggiare l'uso dell'intelligenza artificiale nel settore pubblico e privato, governare le trasformazioni socioeconomiche che essa è destinata a causare e costruire un sistema di garanzie etiche e legali per lo sviluppo di un'intelligenza artificiale affidabile e antropocentrica[71].

Contestualmente fu adottato il *Coordinated Plan on Artificial Intelligence*, all'esito di una lunga consultazione tra la Commissione e gli Stati membri. Esso fu sottoscritto

[70] Gli Stati Uniti hanno adottato una prima *American AI Initiative* nel febbraio 2019. Il piano è stato riorganizzato con il *National Artificial Intelligence Initiative Act*, in vigore dal gennaio 2021 e approvato con il voto trasversale di democratici e repubblicani. L'Unione Europea nel 2018 ha pubblicato la Comunicazione della Commissione al Parlamento Europeo e al Consiglio *L'intelligenza artificiale per l'Europa*. Il Regno Unito, nel 2018 ha lanciato il proprio piano strategico per l'intelligenza artificiale, con la pubblicazione dell'*AI Sector Deal* seguito, nel settembre 2021, dalla *National AI strategy*. Nel 2017 la Cina ha lanciato il *New Generation Artificial Intelligence Development Plan*, con obiettivi da attuare entro il 2030. Analogamente il Giappone nel 2017 ha presentato l'*Artificial Intelligence Technology Strategy* aggiornata nel 2019 e la Corea del Sud nel 2017 ha elaborato il *Mid to Long Term Master Plan in Preparation for the Intelligent Information Society: Managing the Fourth Industrial Revolution*.

[71] Cfr § 3 della Comunicazione della Commissione al Parlamento Europeo e al Consiglio *L'intelligenza artificiale per l'Europa. La strada da seguire: un'iniziativa dell'Ue per l'IA*, 25 aprile 2018, COM/2018/237

anche dalla Norvegia e dalla Svizzera e definì una serie di azioni comuni, necessarie per incrementare gli investimenti nel settore dell'intelligenza artificiale, per favorire la disponibilità di dati con cui sviluppare tecnologie intelligenti e per aumentare il numero di persone con le necessarie competenze tecnologiche. Il documento precisava, altresì, che i settori nei quali l'investimento sull'intelligenza artificiale era prioritario, erano la salute, la mobilità, la sicurezza, l'energia, la produzione industriale e i servizi finanziari[72].

Dall'analisi della strategia e del piano si evince che la Commissione Europea intendeva promuovere lo sviluppo di un'intelligenza artificiale affidabile e antropocentrica. In particolare, nel *Coordinated Plan on Artificial Intelligence* si precisa testualmente che: «*the ambition is for Europe to become the world-leading region for developing and deploying cutting-edge, ethical and secure AI, promoting a human-centric approach in the global context*»[73].

Per realizzare gli obiettivi individuati nei documenti programmatici la Commissione Europea ha istituito un organo consultivo: l'*High-Level Expert Group*. Il gruppo, costituito da cinquantadue esperti del mondo delle imprese, della ricerca, dell'università e della pubblica amministrazione, è stato incaricato di avviare un dialogo sull'intelligenza artificiale con tutti gli stakeholder e di proporre analisi e indicazioni per le politiche europee. In particolare, il gruppo nel 2019 ha elaborato le *Ethics Guidelines for Trustworthy AI*[74] dove è precisato che un'intelligenza artificiale affidabile si basa su tre componenti che dovrebbero essere presenti durante l'intero ciclo di vita del sistema. La prima componente è la legalità; pertanto, i sistemi di intelligenza artificiale devono essere conformi a tutte le leggi e ai regolamenti applicabili. In secondo luogo, essi devono essere etici e, dunque, assicurare l'adesione a principi e valori. Inoltre, i sistemi di intelligenza artificiale devono essere robusti dal punto di vista tecnico e sociale poiché, anche con le migliori intenzioni, essi possono causare danni non intenzionali. Le tre componenti operano armonicamente e si sovrappongono. Le linee guida in esame

[72] Comunicazione della Commissione al Parlamento Europeo, al Consiglio, al Comitato Economico e Sociale Europeo e al Comitato delle Regioni, *Coordinated plan on artificial intelligence*, 7 dicembre 2018, COM /2018/795.

[73] *Coordinated plan on artificial intelligence*, cit., p.1 «*l'ambizione è che l'Europa diventi la regione leader a livello mondiale nello sviluppo e nella diffusione di un'intelligenza artificiale all'avanguardia, etica e sicura, promuovendo un approccio incentrato sull'uomo nel contesto globale*».

[74] HLEG, *Ethics Guidelines for Trustworthy AI (Orientamenti etici per un IA affidabile),* https://op.europa.eu/

non si occupano della legalità, ma intendono offrire orientamenti per promuovere e garantire l'eticità e la robustezza dell'intelligenza artificiale.

Per attuarle concretamente, il gruppo ha adottato le *Policy and Investment Recommendations for Trustworthy AI*[75] e la *Assessment List for Trustworthy AI*[76], con le quali sono state fornite indicazioni pratiche ed è stato introdotto un procedimento di verifica delle caratteristiche di un sistema intelligente affidabile.

Negli anni successivi l'UE, a dimostrazione dell'interesse per il tema dell'intelligenza artificiale, ha elaborato altri documenti[77] finché il 21 aprile 2021 la Commissione Europea ha avviato l'iter legislativo ordinario per l'approvazione di un Regolamento per disciplinare l'intelligenza artificiale con norme giuridicamente vincolanti. Pertanto, è stata pubblicata una proposta di Regolamento indirizzata al Parlamento europeo e al Consiglio «*che stabilisce regole armonizzate sull'intelligenza artificiale (Legge sull'intelligenza artificiale) e modifica alcuni atti legislativi dell'Unione*».[78]

Al termine di un lungo percorso il Regolamento 2024/1689 è stato approvato dal Parlamento Europeo il 13 marzo 2024 e dal Consiglio dell'UE, il 21 maggio 2024. Esso, noto come "*AI Act*", è stato pubblicato nella Gazzetta Ufficiale dell'UE il 12 luglio 2024 ed è entrato in vigore il 2 agosto 2024[79]. Considerata la complessità del testo la

---

[75] HLEG, *Policy and investment recommendations for trustworthy Artificial Intelligence*, (*Raccomandazioni su politiche e investimenti per un'intelligenza artificiale affidabile*), 8 aprile 2019 https://digital-strategy.ec.europa.eu/

[76] HLEG, *Assessment List for Trustworthy AI* (*Elenco di valutazione per l'intelligenza artificiale affidabile*) 17 luglio 2020 https://digital-strategy.ec.europa.eu/

[77] COMMISSIONE EUROPEA, *Libro bianco sull'intelligenza artificiale - Un approccio europeo all'eccellenza e alla fiducia* (COM (2020) 65 final); Comunicazione della Commissione *Plasmare il futuro digitale dell'Europa* (COM (2020) 67 final); Risoluzione del Parlamento europeo del 20 ottobre 2020 recante *Raccomandazioni alla Commissione concernenti il quadro relativo agli aspetti etici dell'intelligenza artificiale, della robotica e delle tecnologie correlate.* 2020/2012(INL); Risoluzione del Parlamento europeo del 20 ottobre 2020 recante *Raccomandazioni alla Commissione su un regime di responsabilità civile per l'intelligenza artificiale*, 2020/2014(INL); Risoluzione del Parlamento europeo del 20 ottobre 2020 sui *Diritti di proprietà intellettuale per lo sviluppo di tecnologie di intelligenza artificiale*,2020/2015(INI); *Progetto di relazione del Parlamento europeo sull'intelligenza artificiale nel diritto penale e il suo utilizzo da parte delle autorità di polizia e giudiziarie in ambito penale*, 2020/2016(INI); *Progetto di relazione del Parlamento europeo sull'intelligenza artificiale nell'istruzione, nella cultura e nel settore audiovisivo*, 2020/2017(INI).

[78] COMMISSIONE EUROPEA, *Proposta di Regolamento al Parlamento Europeo e al Consiglio che stabilisce regole armonizzate sull'intelligenza artificiale (Legge sull'intelligenza artificiale) e modifica alcuni atti legislativi dell'Unione*, 21 aprile 2021, COM (2021) 206. Per un commento alla proposta, G. CONTISSA, F. GALLI, F. GODANO, G. SARTOR, Il regolamento europeo sull'intelligenza artificiale. analisi informatico -giuridica, in *i-lex Rivista di Scienze Giuridiche, Scienze Cognitive ed Intelligenza Artificiale*, 2021, 2.

[79] Regolamento (UE) 2024/1689 del Parlamento europeo e del Consiglio, del 13 giugno 2024, che stabilisce regole armonizzate sull'intelligenza artificiale e modifica i regolamenti (CE) n, 300/2008, (UE) n, 167/2013, (UE) n, 168/2013, (UE) 2018/858, (UE) 2018/1139 e (UE) 2019/2144 e le direttive

Commissione Europea ha, altresì, pubblicato il documento "*Intelligenza artificiale — Domande e risposte*" per fornire un quadro sinottico esplicativo della normativa[80].

L'*AI Act* è la prima legge globale che disciplina l'intelligenza artificiale con norme giuridicamente vincolanti e si basa sul presupposto che tale tecnologia è in rapida evoluzione e contribuisce al conseguimento di un'ampia gamma di benefici a livello economico, ambientale e sociale. In particolare, i sistemi di intelligenza artificiale consentono di migliorare le previsioni, di ottimizzare le operazioni e l'assegnazione delle risorse e di personalizzare le soluzioni digitali disponibili per i singoli e le organizzazioni. Tuttavia, l'impiego di tali sistemi può pregiudicare gli interessi pubblici e i diritti fondamentali tutelati dal diritto dell'UE.

Il Regolamento, dunque, stabilisce che l'intelligenza artificiale deve essere una tecnologia antropocentrica, impiegata per le persone, con il fine ultimo di migliorarne il benessere. Pertanto, prevede norme conformi ai valori dell'UE che garantiscano un livello elevato di protezione della salute, della sicurezza e dei diritti fondamentali sanciti dalla Carta dei diritti fondamentali dell'UE, compresi la democrazia, lo Stato di diritto e la protezione dell'ambiente, e che proteggano contro gli effetti nocivi dei sistemi di intelligenza artificiale nell'UE.

La disciplina unionale promuove, altresì, l'innovazione e la libera circolazione transfrontaliera di beni e servizi basati sull'intelligenza artificiale, impedendo agli Stati membri di imporre unilateralmente restrizioni allo sviluppo, alla commercializzazione e all'uso dei relativi sistemi.

La regolamentazione dei sistemi di intelligenza artificiale in ambito UE, dunque, deve essere armonizzata per evitare che nei singoli Stati membri siano adottate normative divergenti che possano determinare una frammentazione del mercato interno e diminuire la certezza del diritto per gli operatori che li sviluppano, li importano o li utilizzano.

2014/90/UE, (UE) 2016/797 e (UE) 2020/1828 (regolamento sull'intelligenza artificiale), https://eur-lex.europa.eu/. C.IURILLI, *Il diritto naturale come limite e contenuto dell'intelligenza artificiale. Prime riflessioni sul nuovo Regolamento Europeo "AI Act"*.in *Judicium* 24.6.2024; V. TEVERE, *L'UE adotta il primo regolamento sull'intelligenza artificiale: una svolta normativa importante*, in *Lo Stato civile italiano* 2024 ,120 ,5/6 ,p.84 ; G.NATALE, A.M.LAZZE', *AI Act, il regolamento europeo sull'intelligenza artificiale: punti di forza e punti di debolezza*, in *Rassegna avvocatura dello Stato*, 2023, 75, 3,p. 123 ; R.VIOLA, L.DE BIASE, *La legge dell'intelligenza artificiale*, Il sole 24 ore Milano, 2024.

[80] COMMISSIONE EUROPEA, *Intelligenza artificiale — Domande e risposte*", https://ec.europa.eu/

La decisione di utilizzare uno strumento forte come il Regolamento è legata alla strategia di influenzare anche gli altri Stati, attraverso il fenomeno noto come *Brussels effect*[81]. Tale fenomeno consiste nella capacità dell'UE di regolare unilateralmente i mercati globali e imporre standard normativi assistiti da sanzione severe anche fuori dal territorio europeo, facendo leva sulla forza che il mercato Europeo esercita sulle imprese multinazionali che operano, oltre che sul proprio territorio, anche sul mercato europeo. La grandezza e la ricchezza del mercato dei consumatori dell'UE, fanno sì che le aziende multinazionali che operano nel settore dell'High Tech, quali ad esempio Amazon, Apple, Facebook, non possano permettersi di non operare al suo interno. Pertanto, per accedere ad un mercato appetibile, tali imprese devono adeguare la propria condotta agli standard dell'UE spesso tra i più rigorosi a livello globale.

In base all'art.3 dell'*AI Act*, si definisce "*sistema di intelligenza artificiale"*, «*un sistema automatizzato progettato per funzionare con livelli di autonomia variabili e che può presentare adattabilità dopo la diffusione e che, per obiettivi espliciti o impliciti, deduce dall'input che riceve come generare output quali previsioni, contenuti, raccomandazioni o decisioni che possono influenzare ambienti fisici o virtuali*».

La definizione è molto ampia e comprende i sistemi in cui l'intelligenza artificiale opera in piena autonomia e con input umani, i sistemi che possono adattarsi come conseguenza delle informazioni che vengono loro fornite e i sistemi che, in base alle informazioni ricevute, imparano a generare output.

Essa si differenzia dalle definizioni presenti in altri documenti[82] in cui, invece, il "*sistema di intelligenza artificiale*" era descritto come «*un sistema basato su software o integrato in dispositivi hardware che mostra un comportamento che simula l'intelligenza, tra l'altro raccogliendo e trattando dati, analizzando e interpretando il proprio ambiente e intraprendendo azioni, con un certo grado di autonomia, per raggiungere obiettivi specifici*». L'*AI Act*, infatti, intende respingere l'idea che l'intelligenza artificiale possa costituire la simulazione dell'intelligenza umana.

Il Regolamento adotta un approccio basato sulla gestione del rischio e, quindi, classifica i sistemi di intelligenza artificiale in base al rapporto tra la probabilità che essi

[81] L'espressione *Brussels Effect* è stata coniata nel 2020 dalla Professoressa della Columbia Law School A. BRADFORD, *The Brussels Effect. How the European Union Rules the World*, Oxford University Press, Oxford, 2020, p. 18.

[82] Risoluzione del Parlamento europeo del 20 ottobre 2020 recante raccomandazioni alla Commissione su un regime di responsabilità civile per l'intelligenza artificiale, 2020/2014(INL).

provochino un danno alla salute, alla sicurezza e ai diritti fondamentali, e la gravità del danno stesso, delineando diversi regimi giuridici.

Pertanto, in primo luogo, considerata la potenziale lesività della dignità della persona, l'art.5 vieta l'uso dei sistemi che manipolano la cognizione e il comportamento individuale, che si basano sulla raccolta casuale di dati di riconoscimento facciale da Internet o attraverso le telecamere a circuito chiuso, che effettuano il riconoscimento delle emozioni nei luoghi di lavoro e nei contesti educativi, che attribuiscono un punteggio sociale, oppure che operano il trattamento di dati biometrici per l'inferenza di dati personali sensibili, come l'orientamento sessuale o le credenze religiose.

In secondo luogo, il Regolamento detta norme specifiche[83] per l'utilizzo di alcuni sistemi definiti "*ad alto rischio*" che presentano un rischio significativo di danno per la salute, la sicurezza o i diritti fondamentali delle persone fisiche. Essi, come stabilisce l'art.6 comma 1, includono i sistemi di intelligenza artificiale impiegati come componente di sicurezza di prodotti o che sono essi stessi prodotti, sottoposti alla disciplina delle leggi europee di armonizzazione elencate nell'Allegato I[84]. Tali prodotti, come stabilito dalle leggi di armonizzazione, prima di essere messi in commercio, devono essere sottoposti a valutazione di conformità da parte di terzi.

In base all'art.6 comma 2, sono ad alto rischio anche i sistemi di intelligenza artificiale elencati nell'Allegato III che la Commissione può modificare aggiungendo ulteriori sistemi oppure rimuovendo quelli ivi indicati, ove il rischio venga meno.

In tale elenco[85], tra gli altri, si fa riferimento anche ai sistemi di intelligenza artificiale utilizzati nel settore dell'amministrazione della giustizia e, in particolare, ai «*sistemi*

[83] Cfr.Artt.6-49.

[84] La legislazione di armonizzazione (cioè, valida per tutti i paesi UE) consiste nella normativa comunitaria elaborata dalle organizzazioni europee di normazione (OEN) per armonizzare le condizioni di commercializzazione dei prodotti all'interno della UE. Le norme armonizzate stabiliscono specifiche tecniche ritenute adatte o sufficienti per conformarsi ai requisiti tecnici previsti dalla legislazione dell'UE. Nel maggio 2023 la Commissione europea ha incaricato le organizzazioni europee di normazione CEN e CENELEC di elaborare e pubblicare entro la fine di aprile 2025 le norme specifiche per individuare i requisiti dei sistemi di intelligenza artificiale ad alto rischio. La Commissione valuterà ed eventualmente approverà tali norme, che saranno pubblicate nella Gazzetta ufficiale dell'UE. Una volta pubblicate, norme concederanno una "*presunzione di conformità*" al Regolamento, ai sistemi di intelligenza artificiale sviluppati in conformità degli stessi.

[85] L'allegato III indica i seguenti settori: infrastrutture critiche (ad es. i trasporti); formazione educativa o professionale (accesso all'istruzione e al percorso professionale, es. punteggio degli esami); identificazione e categorizzazione biometrica delle persone fisiche; occupazione, gestione dei lavoratori (ad es. software di selezione dei CV per le procedure di assunzione); accesso e fruizione di servizi privati e pubblici essenziali; gestione della migrazione, dell'asilo e del controllo delle frontiere (ad es., la verifica

destinati a essere usati da un'autorità giudiziaria o per suo conto per assistere un'autorità giudiziaria nella ricerca e nell'interpretazione dei fatti e del diritto e nell'applicazione della legge a una serie concreta di fatti, o a essere utilizzati in modo analogo nella risoluzione alternativa delle controversie» .

Il Regolamento, dunque, assegna all'intelligenza artificiale un ruolo accessorio ed eventuale rispetto all'intelligenza del giudice, nell'attività giurisdizionale.

Il "*considerando*" n. 61 dell'*AI Act* spiega, altresì, che «*Alcuni sistemi di IA destinati all'amministrazione della giustizia e ai processi democratici dovrebbero essere classificati come sistemi ad alto rischio, in considerazione del loro impatto potenzialmente significativo sulla democrazia, sullo Stato di diritto, sulle libertà individuali e sul diritto a un ricorso effettivo e a un giudice imparziale. È in particolare opportuno, al fine di far fronte ai rischi di potenziali distorsioni, errori e opacità, classificare come ad alto rischio i sistemi di IA destinati a essere utilizzati da un'autorità giudiziaria o per suo conto per assistere le autorità giudiziarie nelle attività di ricerca e interpretazione dei fatti e del diritto e nell'applicazione della legge a una serie concreta di fatti. Anche i sistemi di IA destinati a essere utilizzati dagli organismi di risoluzione alternativa delle controversie a tali fini dovrebbero essere considerati ad alto rischio quando gli esiti dei procedimenti di risoluzione alternativa delle controversie producono effetti giuridici per le parti. L'utilizzo di strumenti di IA può fornire sostegno al potere decisionale dei giudici o all'indipendenza del potere giudiziario, ma non dovrebbe sostituirlo: il processo decisionale finale deve rimanere un'attività a guida umana. Non è tuttavia opportuno estendere la classificazione dei sistemi di IA come ad alto rischio ai sistemi di IA destinati ad attività amministrative puramente accessorie, che non incidono sull'effettiva amministrazione della giustizia nei singoli casi, quali l'anonimizzazione o la pseudonimizzazione di decisioni, documenti o dati giudiziari, la comunicazione tra il personale, i compiti amministrativi.*».

I sistemi di intelligenza artificiale indicati dall'art.6 comma 2 del Regolamento sono considerati ad alto rischio a meno che il fornitore, attraverso una valutazione di impatto, non dimostri che, per le proprie peculiarità tecniche, non lo sono.

---

dell'autenticità dei documenti di viaggio); applicazione della legge; amministrazione della giustizia e processi democratici.

I sistemi in esame sono sottoposti ad un regime che impone al fornitore la registrazione delle loro attività, per garantirne la tracciabilità e la creazione nonché l'implementazione di un modello procedimentale e documentale di gestione del rischio, per identificare i rischi e adottare azioni di mitigazione durante l'intero ciclo di vita. La loro immissione sul mercato, dunque, richiede una preventiva autovalutazione di conformità ai requisiti richiesti dal Regolamento e l'adozione di appropriate pratiche di gestione e governance risultante da documentazione tecnica da tenere costantemente aggiornata.

Il Regolamento stabilisce, altresì, che i sistemi di intelligenza artificiale ad alto rischio siano progettati e sviluppati in modo da garantire che il loro funzionamento sia sufficientemente trasparente e che gli utenti possano interpretare i risultati del sistema e utilizzarli in modo appropriato.

Inoltre, deve essere consentita una supervisione efficace da parte di persone fisiche durante il periodo in cui il sistema di intelligenza artificiale è in uso, anche con strumenti di interfaccia uomo-macchina appropriati e deve essere garantito un livello adeguato di accuratezza, robustezza e cybersicurezza, durante tutto il ciclo di vita.

I sistemi di intelligenza artificiale ad alto rischio che producono decisioni automatizzate devono essere in grado di fornire alle persone una spiegazione della procedura decisionale e degli elementi principali della decisione presa.

La terza tipologia di sistemi di intelligenza artificiale disciplinata dal Regolamento comprende quelli che, come stabilito dall'art.3, sono «*basati su un modello di IA per finalità generali e che ha la capacità di perseguire varie finalità, sia per uso diretto che per integrazione in altri sistemi di IA*». Anche i modelli di intelligenza artificiale per finalità generali possono comportare dei rischi sistemici[86] ove siano formati utilizzando una potenza di calcolo totale di oltre 10 ^ 25 FLOPS[87] per il loro addestramento. Infatti, le capacità dei modelli al di sopra di tale soglia non sono ancora sufficientemente note. La Commissione Europea può aggiornare o integrare la soglia del rischio sistemico, tenendo conto dei progressi compiuti nella misurazione obiettiva delle capacità del modello e degli sviluppi della potenza di calcolo necessaria per un determinato livello di prestazioni.

[86] Cfr.Art.51

[87] Flop (*floating point operations*) è un indicatore delle capacità del modello.

Il Regolamento[88] pone a carico dei fornitori dei modelli per finalità generali, degli obblighi di trasparenza per garantire che chiunque ne faccia uso, anche per creare altri sistemi di intelligenza artificiale, abbia accesso a tutte le informazioni necessarie. In particolare, devono essere indicate le modalità seguite per sviluppare il sistema, le attività svolte e il consumo energetico stimato.

I fornitori dei modelli in esame devono, altresì, disporre di politiche volte a garantire il rispetto del diritto d'autore e redigere e rendere disponibile al pubblico un riassunto sufficientemente dettagliato dei contenuti utilizzati per l'addestramento, secondo un modello fornito dall'Ufficio per l'IA.

Nel caso in cui il modello per finalità generali abbia un rischio sistemico i fornitori devono eseguire la valutazione, mitigare i possibili rischi, tenere traccia, documentare e riferire le informazioni pertinenti sugli incidenti gravi e le possibili misure correttive per affrontarli e garantire un livello adeguato di protezione della cybersecurity[89]. Essi, sono invitati a collaborare con l'Ufficio per l'IA e altri portatori di interessi per elaborare un codice di buone pratiche che specifichi le norme e garantisca, in tal modo, lo sviluppo sicuro e responsabile dei loro modelli.

Il Regolamento[90] stabilisce che tutti i sistemi di intelligenza artificiale, anche se non sono rischiosi, impongono obblighi di trasparenza di base, per garantire un livello minimo di chiarezza e comprensione, informando le persone che interagiscono con loro. I fornitori di tali sistemi possono scegliere di applicare, su base volontaria, i requisiti per un'intelligenza artificiale affidabile e aderire a codici di condotta volontari.

Le norme dell'*AI Act* si applicano ai fornitori (*provider)* e agli utilizzatori (*deployers*).

I primi sono persone fisiche o giuridiche, autorità pubbliche, agenzie o altri organismi che sviluppano o fanno sviluppare un sistema o un modello per finalità generali di intelligenza artificiale e li immettono sul mercato con il proprio nome o marchio, a titolo oneroso o gratuito.

I fornitori, dunque, sono le entità giuridiche soggette agli obblighi più rilevanti sanciti dal Regolamento.

Gli utilizzatori, invece, sono persone fisiche o giuridiche, autorità pubbliche, agenzie o altri organismi che utilizzano un sistema di intelligenza artificiale sotto la propria

[88] Cfr. Art.53
[89] Cfr. Art.55
[90] Cfr. Art.50

autorità, tranne nel caso in cui esso sia utilizzato nel corso di un'attività personale non professionale.

In base al Regolamento gli utilizzatori devono adottare misure tecniche e organizzative adeguate per garantire l'utilizzo dei sistemi in conformità alle istruzioni d'uso che li accompagnano, assegnare la supervisione umana a persone fisiche che abbiano la competenza, la formazione e l'autorità necessarie, nonché il supporto necessario, monitorare il funzionamento dei sistemi e adempiere agli obblighi di informazione prima della messa in funzione del sistema, nonché agli obblighi imposti dalla normativa a tutela della riservatezza dei dati.

Con specifico riferimento ai sistemi di intelligenza artificiale ad alto rischio che sono utilizzati per la valutazione del merito creditizio e per la valutazione del rischio e la determinazione dei prezzi in relazione alle persone fisiche nel caso dell'assicurazione sulla vita e sulla salute, gli implementatori devono eseguire anche la valutazione d'impatto sui diritti fondamentali (FRIA) che deve essere notificata all'autorità di vigilanza del mercato con i relativi risultati.

Il Regolamento si applica a tutti i sistemi di intelligenza artificiale immessi sul mercato dell'UE o il cui utilizzo abbia un impatto sulle persone situate nell'UE.

Tuttavia, esso non trova applicazione per le attività di ricerca, sviluppo e prototipazione svolte prima dell'immissione sul mercato di un sistema di intelligenza artificiale e per i sistemi progettati esclusivamente per finalità militari, di difesa o di sicurezza nazionale, indipendentemente dal tipo di entità che svolge tali attività.

Dal punto di vista della governance l'*AI Act* ha istituito un sistema a due livelli, in cui le autorità nazionali sono responsabili della supervisione e dell'applicazione delle norme per i sistemi di intelligenza artificiale[91]. Il livello dell'UE è rappresentato dall'Ufficio per l'intelligenza artificiale presso la Commissione europea che svolge un ruolo chiave a sostegno delle autorità nazionali ed è responsabile della regolamentazione dei modelli

[91] Cfr. art.59 Alle autorità nazionali sono conferite due funzioni: la funzione di vigilanza e quella di notificazione. La prima è finalizzata a controllare che l'*AI Act* sia rispettato da parte dei produttori e distributori di sistemi di intelligenza artificiale; la funzione di notificazione, invece, consiste nella verifica di regolarità delle attività di certificazioni rilasciate da soggetti terzi a chi crea sistemi di intelligenza artificiale che rientrano nelle categorie ad alto rischio. Il regolamento prevede che le Autorità nazionali siano costituite da soggetti con una forte specializzazione tecnica e siano istituite attraverso fonte primaria nazionale coordinata con l'ordinamento europeo.

di intelligenza artificiale per finalità generali, può richiedere informazioni e misure ai fornitori di modelli e applicare sanzioni.

Il Regolamento prevede l'istituzione di un Comitato europeo per l'intelligenza artificiale composto da rappresentanti degli Stati membri, con sottogruppi specializzati per le autorità nazionali di regolamentazione e altre autorità competenti.

Il Regolamento prevede, altresì, la creazione di un panel scientifico e di un forum consultivo per integrare le prospettive delle diverse parti interessate, in modo che la normativa sia sempre aggiornata rispetto agli sviluppi della tecnologia.

L'*AI Act* prevede un sistema di sanzioni che si applicano, altresì, alle istituzioni, alle agenzie o agli organi dell'UE. Anche il Garante europeo della protezione dei dati ha il potere di imporre sanzioni pecuniarie in caso di inosservanza. Il sistema sanzionatorio è basato sul fatturato globale delle aziende o su un importo predeterminato, a seconda di quale sia più alto. Sono previste eccezioni per le aziende più piccole, con sanzioni limitate per le PMI e le startup.

L'*AI Act* è entrato in vigore il 2 agosto 2024 nella parte relativa alle norme generali. A partire dal 2 febbraio 2025 entrerà in vigore il divieto di utilizzo dei sistemi con rischio inaccettabile sancito dall'art.5.

Pertanto, a partire da tale data è vietata l'immissione sul mercato, la messa in servizio o l'uso di sistemi di IA che distorcano materialmente il comportamento di una persona o di un gruppo di persone utilizzando tecniche subliminali manipolative, inducendole a prendere una decisione che non avrebbero altrimenti preso, in un modo che provochi o possa ragionevolmente provocare un danno significativo. Inoltre, sono vietati i sistemi che distorcono materialmente il comportamento di una persona o di un gruppo di persone, sfruttandone le vulnerabilità, quali ad esempio età, disabilità o situazione socioeconomica, in un modo che provochi o possa ragionevolmente provocare un danno significativo. Sono, altresì, vietati sistemi di social scoring fondati sulla valutazione o classificazione di persone o gruppi di persone sulla base del loro comportamento sociale o di caratteristiche personali o della personalità note, inferite o previste che determinano trattamenti pregiudizievoli o sfavorevoli ingiustificati o sproporzionati.

Il divieto di immissione sul mercato si estende ai sistemi che valutano il rischio che una persona commetta un reato, unicamente sulla base della sua profilazione o della valutazione dei tratti e delle caratteristiche della sua personalità; tale divieto non si

applica ai sistemi di IA utilizzati a sostegno della valutazione umana del coinvolgimento di una persona in un'attività criminosa, che si basa già su fatti oggettivi e verificabili direttamente connessi a un'attività criminosa.

Sono vietati anche i sistemi che consentono di creare o ampliare le banche dati di riconoscimento facciale mediante scraping non mirato di immagini facciali da internet o da filmati di telecamere a circuito chiuso.

Il divieto dell'art.5 si applica, altresì, ai sistemi che inferiscono le emozioni di una persona fisica nell'ambito del luogo di lavoro e degli istituti di istruzione, tranne il caso in cui il sistema di IA sia destinato ad essere impiegato per motivi medici o di sicurezza.

Il Regolamento vieta i sistemi che inferiscono razza, opinioni politiche, appartenenza sindacale, convinzioni religiose o filosofiche, vita sessuale o orientamento sessuale, classificando individualmente le persone fisiche sulla base dei loro dati biometrici. È, invece, consentita la classificazione di set di dati biometrici acquisiti legalmente oppure utilizzati per le attività connesse alla tutela dell'ordine pubblico e della sicurezza nazionale.

Infine, sono vietati i sistemi che permettono l'identificazione biometrica remota "*in tempo reale*" in spazi accessibili al pubblico a meno che non siano finalizzati alla ricerca di vittime di rapimento, tratta, sfruttamento sessuale e persone scomparse, prevenzione di minacce imminenti alla vita, all'incolumità fisica o relative ad attacchi terroristici oppure all'identificazione di sospetti in indagini penali per reati gravi punibili con almeno quattro anni di reclusione.

Le aziende, dunque, devono analizzare quali sistemi e modelli di IA stanno già utilizzando o prevedono di utilizzare nel prossimo futuro, valutare il relativo rischio e sviluppare strategie di risposta nel caso emergano aree di potenziale non conformità.

Il 2 agosto 2025, invece, entreranno in vigore le regole di governance e gli obblighi per i modelli di intelligenza artificiale per finalità generali.

Gli adempimenti relativi ai sistemi ad alto rischio saranno obbligatori a partire dal 2 agosto 2026 per i sistemi indicati nell'allegato III.

Infine, il 2 agosto 2027 entreranno in vigore le disposizioni applicabili ai prodotti armonizzati indicati nell'allegato I.

L'applicazione del Regolamento prevede la pubblicazione di una serie di provvedimenti attuativi e, in particolare, degli orientamenti sugli usi di IA vietati e delle Linee guida

per la classificazione ad alto rischio. Inoltre, la Commissione ha affidato alle organizzazioni europee di standardizzazione CEN e CENELEC, l'attività di sviluppo degli standard tecnici. Le norme dovranno essere sviluppate e pubblicate entro la fine di aprile 2025. Infine, saranno adottate linee guida per fornire ulteriori indicazioni ai fornitori e agli operatori per i sistemi ad alto rischio.

All'*AI Act* è strettamente connessa la proposta di *Direttiva sulla responsabilità civile da intelligenza artificiale*[92] che integra e aggiorna il quadro dell'UE in materia di responsabilità civile, introducendo per la prima volta norme specifiche per i danni causati dai sistemi di IA.

Le nuove norme consentiranno ai soggetti danneggiati dalla tecnologia di IA di accedere al risarcimento come se avessero subito danni in qualsiasi altra circostanza[93].

Le norme nazionali vigenti in materia di responsabilità civile non sono applicabili nel caso di danni causati da prodotti e servizi basati sull'IA. Infatti, nelle azioni di responsabilità civile il danneggiato deve identificare chi citare in giudizio e fornire la prova dell'evento, del danno e del nesso di causalità tra i due. Quando si tratta di IA, l'onere della prova sarebbe insostenibile. Pertanto, la Direttiva introduce una presunzione di causalità. Se i danneggiati possono dimostrare che per negligenza non sono stati rispettati determinati obblighi e che è ragionevolmente probabile un nesso di causalità tra l'evento e il danno, il giudice può applicare la presunzione, che può essere superata con la prova che il danno è stato provocato da una causa diversa. Inoltre, ai danneggiati è riconosciuta la possibilità di utilizzare lo strumento dell'ordine giudiziale di esibizione per conseguire informazioni sui sistemi di IA ad alto rischio, come quelli impiegati in ambito medico o di sicurezza, in modo da poter identificare la persona che potrebbe essere ritenuta responsabile e scoprire cosa non ha funzionato, nel rispetto dei segreti commerciali. Se un'azienda non ottempera a un ordine di esibizione, si configura una presunzione di colpa, che alleggerisce ulteriormente l'onere probatorio per il danneggiato.

[92] Proposta di DIRETTIVA DEL PARLAMENTO EUROPEO E DEL CONSIGLIO relativa all'adeguamento delle norme in materia di responsabilità civile extracontrattuale all'intelligenza artificiale (direttiva sulla responsabilità da intelligenza artificiale) COM/2022/496

[93] R. PANETTA, *Direttiva UE su responsabilità per danni da intelligenza artificiale: passo necessario per l'Europa*, in *AgendaDigitale*, su https://www.agendadigitale.eu, 12 Marzo 2024.

Anche la Direttiva in esame si basa sulla distinzione presente nell'*AI Act* tra sistemi IA a basso rischio e ad alto rischio, con un livello di regolamentazione proporzionato al grado di rischio associato al loro impiego.

È, altresì, prevista la possibilità di una responsabilità oggettiva per le aziende che operano nel settore dell'IA che possono essere considerate responsabili dei danni causati dai loro sistemi, indipendentemente dalla colpa.

Tale regime, integrato dall'obbligo di assicurazione per i produttori di sistemi IA ad alto rischio, mira a creare un quadro normativo solido che incentivi le imprese a sviluppare nuove tecnologie senza dover affrontare rischi eccessivi in caso di malfunzionamenti o danni.

La Direttiva sulla responsabilità civile si connette, non solo con l'*AI Act*, ma anche con l'aggiornamento della Direttiva sulla responsabilità per prodotti difettosi[94], che include le nuove tecnologie. L'ambito oggettivo di applicazione di tale Direttiva, infatti, è costituito dal "*prodotto*" che include anche software, IA e servizi digitali, riconoscendo che i danni possono derivare, non solo da difetti fisici, ma anche da malfunzionamenti software o decisioni errate prese da sistemi autonomi.

Il legislatore europeo ha previsto una riduzione della soglia per i danni risarcibili, garantendo che anche danni di minore entità, ma potenzialmente frequenti nell'uso quotidiano di tecnologie IA, possano essere oggetto di risarcimento.

Le aziende, in piena coerenza con quanto stabilito nell'*AI Act*, sono sollecitate a sviluppare sistemi IA "*spiegabili*", trasparenti, basati su processi comprensibili dagli esseri umani. La trasparenza, infatti, facilita l'imputazione della responsabilità in caso di danni e promuove una maggiore fiducia del pubblico nell'IA.

La Direttiva in esame affronta anche la questione della responsabilità in scenari in cui i sistemi IA si evolvono nel tempo basandosi sui dati che ricevono, proponendo meccanismi per gestire tale evoluzione, inclusa la possibilità di "*congelamento*" delle versioni dei sistemi IA per scopi di analisi forensi in caso di incidenti.

[94] DIRETTIVA (UE) 2024/2853 DEL PARLAMENTO EUROPEO E DEL CONSIGLI del 23 ottobre 2024 sulla responsabilità per danno da prodotti difettosi, che abroga la direttiva 85/374/CEE del Consiglio.

1.4 La Convenzione quadro sull'intelligenza artificiale e i diritti umani del Consiglio d'Europa del 17 maggio 2024

Il Consiglio d'Europa[95], il 17 maggio 2024, ha adottato *La Convenzione quadro del Consiglio d'Europa sull'intelligenza artificiale e i diritti umani, la democrazia e lo Stato di diritto* che rappresenta il primo trattato internazionale giuridicamente vincolante[96].

La Convenzione giunge all'esito di due anni di lavoro svolto da un organismo intergovernativo, il Comitato sull'intelligenza artificiale (CAI), che riunisce gli Stati membri del Consiglio d'Europa, tra i quali sono compresi gli Stati membri dell'UE e undici Stati non membri[97], nonché rappresentanti del settore privato, della società civile e del mondo accademico, che hanno partecipato in qualità di osservatori.

Essa intende definire un quadro giuridico sostenuto da Stati di diversi continenti uniti da valori comuni, che consenta di trarre vantaggio dall'intelligenza artificiale, riducendo, al contempo, i rischi che questa presenta per i diritti umani, la democrazia e lo Stato di diritto.

Pertanto, il trattato in esame non intende regolare questioni economiche e di mercato e adotta un approccio antropocentrico, basato sul rischio per i diritti umani, connesso alle attività svolte nell'ambito dell'intero ciclo di vita dei sistemi di intelligenza artificiale da parte di attori sia pubblici, che privati.

La Convenzione, come si è detto[98], stabilisce che «*per "sistema di intelligenza artificiale" si intende un sistema basato su una macchina che, per obiettivi espliciti o impliciti, deduce, dagli input che riceve, come generare output quali previsioni, contenuti, raccomandazioni o decisioni che possono influenzare le condizioni fisiche o ambienti virtuali. Diversi sistemi di intelligenza artificiale variano nei loro livelli di autonomia e ad attività dopo l'implementazione*».

---

[95] Il Consiglio d'Europa ha sede a Strasburgo ed è un'organizzazione intergovernativa istituita con il Trattato di Londra del 5 maggio 1949 da 10 Stati fondatori (Belgio, Francia, Lussemburgo, Paesi Bassi, Regno Unito, Danimarca, Norvegia, Irlanda, Italia e Svezia) con l'obiettivo di favorire la creazione di uno spazio democratico e giuridico comune in Europa, garantendo il rispetto dei diritti umani, della democrazia e dello stato di diritto. Oggi ha una dimensione paneuropea, comprendendo 46 Stati. Hanno status di osservatore la Santa Sede, gli Stati Uniti, il Canada, il Giappone e il Messico.

[96] Consiglio d'Europa, *Convenzione quadro sull'intelligenza artificiale e i diritti umani, la democrazia e lo Stato di diritto*, https://rm.coe.int/1680afae3c

[97] Argentina, Australia, Canada, Costa Rica, Giappone, Israele, Messico, Perù, Santa Sede, Stati Uniti d'America e Uruguay

[98] Cfr. supra §1.1

Il Consiglio d'Europa condivide la definizione elaborata dall'OCSE, inserita nella Raccomandazione del Consiglio sull'intelligenza artificiale, aggiornata l'8 novembre 2023 affinché continui a essere tecnicamente precisa e a riflettere gli sviluppi tecnologici. Tale definizione, nella sostanza, corrisponde a quella dell'*AI Act*. Infatti, entrambe individuano le medesime proprietà chiave dei sistemi di intelligenza artificiale: autonomia e adattabilità variabili, capacità di inferenza e di generazione di previsioni, contenuti, raccomandazioni o decisioni che possono influenzare ambienti fisici o virtuali.

La Convenzione, come l'*AI Act* è basata sul rischio. Tuttavia, non individua il proprio ambito di applicazione in specifici modelli, sistemi o pratiche. Essa intende regolamentare le specifiche attività del ciclo di vita dei sistemi di intelligenza artificiale che possono interferire con i diritti umani, la democrazia e lo stato di diritto, svolte sia da soggetti pubblici, che privati. Pertanto, a differenza dell'*AI Act*, analizza le singole attività facenti parte del ciclo di vita dell'intelligenza artificiale e l'impatto che esse possono avere, anche a prescindere dal rischio che il sistema presenta nel suo complesso.

I primi due capitoli del provvedimento sanciscono previsioni e obblighi generali. Il terzo capitolo stabilisce una serie di principi generali da attuare in conformità con gli ordinamenti giuridici nazionali. Essi sono formulati in modo da poter essere adattati a contesti in rapida evoluzione.

In primo luogo, la Convenzione, applicando un approccio antropocentrico, prescrive l'adozione, da parte degli Starti aderenti, di misure per il rispetto della dignità umana e dell'autonomia individuale. Tali misure devono garantire alle persone il controllo sull'utilizzo e l'impatto delle tecnologie di intelligenza artificiale nelle loro vite.

L'art.8 sancisce il principio di trasparenza e di supervisione dei sistemi di intelligenza artificiale prescrivendo l'adozione di adeguati presidi sotto questi due aspetti. Attraverso tali presidi deve essere garantita la comprensione e l'accessibilità dei processi decisionali e del funzionamento generale dei sistemi in esame.

Il principio di accountability e responsabilità impone l'adozione di misure atte a garantire che le organizzazioni, le entità e gli individui impegnati nelle attività lungo tutto il ciclo di vita dei sistemi di intelligenza artificiale rispondano degli impatti negativi sui diritti umani, sulla democrazia o sullo Stato di diritto.

La Convenzione enuncia, altresì, i principi di eguaglianza e non discriminazione, di tutela dei dati personali, di affidabilità, e, infine, di innovazione sicura in ambienti controllati.

Il quarto capitolo è dedicato ai rimedi da applicare in caso di lesione dei diritti umani della democrazia e dello stato di diritto. La Convenzione prevede che ciascuno Stato ad essa aderente applichi i propri regimi normativi alle attività che caratterizzano il ciclo di vita dei sistemi di intelligenza artificiale. Si richiede, altresì, l'adozione e il mantenimento di misure specifiche volte a documentare e rendere disponibili determinate informazioni alle persone interessate, ma anche ad assicurare l'effettiva possibilità di presentare reclamo alle autorità competenti.

Analogamente all'*AI Act* anche nella Convenzione in esame è previsto che i soggetti che interagiscono con i sistemi di intelligenza artificiale debbano esserne resi consapevoli, attraverso adeguati mezzi informativi.

Il quinto capitolo costituito dall'art.16 prevede l'obbligo di identificare, valutare, prevenire e mitigare i rischi e gli impatti potenziali rilevanti per i diritti umani, la democrazia e lo stato di diritto, ex ante e, se necessario, in modo iterativo per tutto il ciclo di vita del sistema di intelligenza artificiale, sviluppando un modello di gestione dei rischi basato su criteri concreti e oggettivi.

La Convenzione impone agli Stati aderenti anche di valutare la necessità dell'adozione di moratorie, divieti o altre misure inibitorie per i sistemi di intelligenza artificiale incompatibili con il rispetto dei diritti umani, della democrazia e dello stato di diritto. Ai singoli Stati è lasciata la libertà di definire i presupposti che legittimano l'adozione di tali misure.

In considerazione delle differenze tra gli ordinamenti giuridici degli Stati che aderiranno alla Convenzione, è stabilito che essi possono scegliere di essere sottoposti direttamente alle disposizioni in essa contenute oppure possono adottare altre misure legislative o amministrative per conformarsi ad essa. L'implementazione della Convenzione deve avvenire nel rispetto dell'uguaglianza, compresa quella di genere, e del divieto di discriminazione, nonché della tutela della privacy.

Gli Stati aderenti non sono tenuti ad applicare la Convenzione al settore della sicurezza nazionale, ma devono comunque assicurare che tali attività siano svolte nel rispetto dei principi democratici. Sono escluse dall'ambito di applicazione della disciplina anche le

attività di difesa nazionale e di ricerca e sviluppo, salvo nei casi in cui il test dei sistemi di intelligenza artificiale potrebbe interferire con i diritti umani, la democrazia e lo stato di diritto.

Gli Stati aderenti possono applicare precedenti accordi o trattati relativi al ciclo di vita dei sistemi di intelligenza artificiale coperti dalla Convenzione, ma devono attenersi agli obiettivi e alle finalità della stessa, senza assumere obblighi in contrasto.

A partire dal 5 settembre 2024 la Convenzione è aperta alla firma, non solo degli Stati membri del Consiglio d'Europa, ma anche dei Paesi terzi che hanno contribuito alla sua elaborazione.

La Convenzione è stata sottoscritta dal Montenegro, Andorra, Georgia, Islanda, Norvegia, Repubblica di Moldova, San Marino e Regno Unito, nonché Israele, Stati Uniti d'America e Unione europea.

1.5. La strategia italiana per l'IA 2024-2026 e il d.d.l. AS1146 presentato dal Governo italiano al Parlamento il 24 maggio 2024.

Come si è detto in precedenza, l'UE, per evitare che nei singoli Stati membri siano adottate normative divergenti che possano determinare una frammentazione del mercato interno e diminuire la certezza del diritto, ha scelto di disciplinare l'intelligenza artificiale attraverso l'adozione di un Regolamento che, come stabilisce l'art. 288 del Trattato sul funzionamento dell'Unione europea (TFUE), è un atto giuridico di portata generale, vincolante in tutti i suoi elementi e direttamente applicabile negli Stati membri. Tuttavia, il Regolamento non limita l'autonomia normativa degli Stati membri in materia di sicurezza nazionale e di ricerca e nei settori che non interferiscono con il mercato interno. Infatti, come si è detto, le regole dell'*AI Act* non si applicano ai sistemi che hanno esclusivamente scopi militari, di difesa o di sicurezza nazionale, nonché alle attività di ricerca, sviluppo e prototipazione che precedono l'immissione sul mercato o a persone che utilizzano l'intelligenza artificiale per motivi non professionali.

Il Regolamento prevede, altresì, l'individuazione di autorità nazionali competenti per la sorveglianza sull'attuazione delle nuove norme e l'applicazione di sanzioni amministrative. Inoltre, esso lascia agli Stati membri la possibilità di prevedere sanzioni di natura penale per garantire la tutela degli interessi che possono essere lesi da un uso illecito dell'intelligenza artificiale.

Il 22 luglio 2024, a pochi giorni dalla pubblicazione dell'*AI Act* sulla Gazzetta Ufficiale dell'UE è stata pubblicata la *Strategia Italiana per l'Intelligenza Artificiale 2024-2026*[99] elaborata da un Comitato di esperti per supportare il Governo nella definizione di una normativa nazionale e delle politiche sull'intelligenza artificiale.

Il Comitato, composto da quattordici membri competenti ed esperti del settore, ha analizzato l'impatto dell'intelligenza artificiale e ha elaborato un piano strategico per guidare lo sviluppo dell'intelligenza artificiale in modo sicuro, etico e inclusivo, massimizzando i benefici e minimizzando i potenziali effetti avversi.

Il documento, dopo aver ricostruito il contesto globale e nazionale entro il quale si sviluppa la ricerca sull'intelligenza artificiale, delinea la visione che deve ispirare le scelte strategiche del Paese. In particolare, si rileva che l'Italia non può limitarsi ad utilizzare i sistemi di intelligenza artificiale ma deve saper sostenere, sia l'attività di ricerca fondazionale, sia quella applicata, creando sinergie con partenariati pubblico-privati. Pertanto, sono necessari investimenti in infrastrutture sempre più efficienti per il calcolo e sempre più dedicate alle applicazioni dell'intelligenza artificiale. Inoltre, è necessario garantire elevati standard qualitativi per la formazione di esperti del settore ai quali devono essere offerte concrete prospettive di crescita personale e professionale[100].

Per realizzare la visione strategica, gli investimenti sull'intelligenza artificiale devono riguardare le molteplici aree e i possibili ambiti di applicazione di tali tecnologie. Tuttavia, una particolare attenzione deve essere riservata agli investimenti in quattro macroaree che sono particolarmente rilevanti per il tessuto produttivo e sociale del Paese: la ricerca, la Pubblica Amministrazione, le imprese e la formazione.

La strategia propone un sistema di monitoraggio della relativa attuazione affidata alla Fondazione per l'Intelligenza Artificiale sotto il diretto controllo della Presidenza del Consiglio dei ministri, dotata di un fondo per supportare la pianificazione e garantire un miglioramento costante del sistema. La Fondazione, con il supporto di un panel di esperti, deve redigere un report annuale, nel quale è mantenuta aggiornata l'analisi di contesto ed è fornita una valutazione complessiva sull'andamento dell'attuazione della Strategia. Oltre che del monitoraggio, la Fondazione sarà responsabile anche dell'attuazione e del coordinamento delle singole iniziative.

[99] *Strategia Italiana per l'Intelligenza Artificiale 2024-2026*, https://www.agid.gov.it/
[100] *Ivi*, pp.6 e 7 https://www.agid.gov.it/

La Strategia, infine, interviene in merito alla modalità di individuazione dell'autorità nazionale per l'intelligenza artificiale che deve essere istituita in attuazione dell'*AI Act*. A tal proposito, occorre rilevare che l'intelligenza artificiale è una tecnologia che riguarda ambiti giuridici già regolati, nei quali operano altre Autorità con i loro differenti indirizzi per specificità di settori. Si pensi, ad esempio, alla tutela della riservatezza dei dati ove opera il Garante privacy, alla tutela dei consumatori e alla disciplina del mercato dei servizi digitali ove opera l'AGCM, al contrasto dei contenuti digitali dannosi di competenza di AGCOM. Vi è, dunque, il rischio di una sovrapposizione di norme e di competenze tra Autorità. Pertanto, l'autorità nazionale per l'intelligenza artificiale deve contribuire a semplificare il contesto e deve, altresì, siglare protocolli e mantenere una stretta collaborazione con l'Agenzia per la Cybersicurezza Nazionale (ACN).

Con riferimento specifico ai quattro settori di intervento, il documento stabilisce che nell'ambito della ricerca scientifica l'Italia deve consolidare la propria competitività sul piano internazionale, aumentando gli investimenti e promuovendo le iniziative co-progettate da partenariati pubblico-privati che coinvolgono le imprese.

Nel settore della pubblica amministrazione, è necessario sfruttare i benefici offerti dall'intelligenza artificiale, per migliorare l'efficienza interna e dei servizi ai cittadini. Pertanto, è essenziale, che siano intraprese azioni multidisciplinari per garantire la privacy, la sicurezza e la corretta gestione dei dati, che siano sviluppati sistemi di intelligenza artificiale interoperabili e che il personale sia adeguatamente formato. Inoltre, il documento prescrive l'adozione di linee guida per promuovere l'impiego di sistemi di intelligenza artificiale nel settore pubblico e nei rapporti con i cittadini e le imprese.

La Strategia, per massimizzare i benefici dell'intelligenza artificiale per il mondo delle imprese, propone, in primo luogo, di valorizzare il ruolo delle imprese ICT italiane nello sviluppo dei relativi sistemi, in collaborazione con università ed enti di ricerca. A tale scopo è necessario semplificare la gestione delle pratiche normative e certificative. Le imprese non direttamente coinvolte nello sviluppo tecnologico, invece, dovrebbero utilizzare l'intelligenza artificiale per accrescere la loro competitività e per affrontare le sfide della sostenibilità ambientale. L'attuazione di tale strategia richiede finanziamenti dedicati che supportino l'adozione e lo sviluppo di soluzioni di intelligenza artificiale

interoperabili, la creazione di laboratori in contesti industriali e la crescita delle start-up operanti nel settore.

Con riferimento alla formazione, infine, il documento propone un piano integrato per rafforzare e diffondere la conoscenza dell'intelligenza artificiale nel sistema educativo, dalle scuole superiori alle università, riservando particolare attenzione ai programmi di dottorato di ricerca. Inoltre, sono necessari programmi di formazione per riqualificare e aggiornare i lavoratori del settore pubblico e privato. La Strategia prevede interventi di alfabetizzazione sull'intelligenza artificiale per tutta la popolazione per evitare che un divario di conoscenza minacci la coesione sociale ed economica nel lungo periodo[101].

Prima che fosse pubblicata la Strategia, il Governo italiano ha deliberato un disegno di legge recante "*Disposizioni e delega al Governo in materia di intelligenza artificiale*" [102] per disciplinare l'uso dell'intelligenza artificiale nei settori demandati dall'*AI Act* all'autonomia normativa degli Stati membri. Esso è stato presentato al Senato il 20 maggio 2024 dove è in corso di esame.

L'obiettivo del disegno di legge, i principi fondamentali e la definizione di "*sistemi di intelligenza artificiale*", in base al testo del disegno di legge, sarebbero mutuati dal Regolamento UE. Tuttavia, il Governo non intende solo accelerare l'introduzione di alcuni dei principi previsti dall'*AI Act* ma provvede anche a disciplinare alcune aree potenzialmente critiche interessate dall'utilizzo dei sistemi di intelligenza artificiale e annunciare lo stanziamento di risorse volte a consentire all'Italia di avere una presenza strategica nel contesto europeo.

Pertanto, oltre ai criteri regolatori, sono previste norme di settore, norme che promuovono l'utilizzo delle nuove tecnologie prevedendo, al contempo, misure in grado di contenere il rischio connesso al loro uso improprio o dannoso, norme che tutelano gli utenti e il diritto di autore e norme che introducono sanzioni penali.

Con riferimento specifico all'attività giudiziaria il disegno di legge in esame[103], stabilisce che l'intelligenza artificiale potrà essere usata solo per attività di supporto quali l'organizzazione e la semplificazione del lavoro oppure per la ricerca

[101] Per realizzare questi obiettivi la Strategia propone di attuare percorsi di apprendimento sull'intelligenza artificiale nelle scuole, di creare programmi di tirocinio, interscambio e visiting in aziende e centri di ricerca, di integrare l'intelligenza artificiale come materia nei corsi di laurea universitari, di sostenere il Dottorato Nazionale in intelligenza artificiale.

[102] D.d.l. AS1146, www.senato.it

[103] Cfr. Artt.14 d.d.l. AS 1146

giurisprudenziale e dottrinale. La norma precisa, altresì, che è sempre riservata al magistrato la decisione sull'interpretazione della legge, la valutazione dei fatti e delle prove e sull'adozione di ogni provvedimento, inclusa la sentenza.

Il disegno di legge riconosce la centralità della cybersicurezza lungo tutto il ciclo di vita dei sistemi di IA. Pertanto, prevede l'implementazione di robuste misure di sicurezza informatica, di protocolli di gestione dei dati e di procedure di accesso rigidamente controllate per proteggere le informazioni sensibili dei cittadini e l'integrità dei sistemi.

Inoltre, è necessario garantire che i sistemi di IA non perpetuino o amplifichino bias esistenti nei vari ambiti di applicazione. Pertanto, occorre procedere con un'attenta selezione e verifica degli algoritmi utilizzati, e con un monitoraggio continuo dei loro output per identificare e correggere eventuali tendenze discriminatorie.

Il legislatore evidenzia anche l'importanza della sostenibilità non solo ambientale, ma anche economica e sociale dell'implementazione di sistemi di IA che richiede un valutare attenta dei costi a lungo termine, considerando non solo l'investimento iniziale ma anche i costi di manutenzione, aggiornamento e formazione continua del personale.

Il disegno di legge attribuisce alla Presidenza del Consiglio dei ministri un ruolo centrale nella definizione e nell'attuazione della Strategia nazionale per l'intelligenza artificiale e individua nell'Agenzia per l'Italia digitale (AgID) e nell'Agenzia per la cybersicurezza nazionale (ACN) le Autorità nazionali per l'intelligenza artificiale[104] previste dall'art.59 dell'*AI Act*, escludendo l'Autorità Garante per la privacy.

Inoltre, si delega al Governo l'adeguamento della normativa nazionale al Regolamento europeo e la definizione della disciplina in caso di uso dell'intelligenza artificiale per finalità illecite.

Dal punto di vista legale e procedurale, il disegno di legge prevede che le cause che hanno ad oggetto il funzionamento di un sistema di IA rientrano nella competenza per materia del tribunale civile e potrebbe rendersi necessario istituire sezioni specializzate all'interno dei tribunali per gestire queste nuove tipologie di controversie. Inoltre, è fondamentale formare i giudici su questioni tecniche relative all'IA per garantire una corretta valutazione di queste cause.

[104] Cfr. Art.18 d.d.l. AS1146

Il provvedimento in esame è stato sottoposto all'Autorità Garante per la privacy che, il 2 agosto 2024[105], ha espresso un parere favorevole, raccomandando, tuttavia, un "*miglior coordinamento sistematico*" con la disciplina in materia di protezione dei dati e chiedendo alcune integrazioni e correzioni per garantire una maggiore tutela dei dati personali dei cittadini.

Tra l'altro, il Garante ha chiesto che nel Capo I sia introdotta una disposizione normativa generale per precisare che i trattamenti di dati personali effettuati attraverso i sistemi di intelligenza artificiale devono rispettare la normativa a tutela della privacy nazionale ed europea.

Inoltre, il Garante ritiene necessario che sia prevista l'adozione di sistemi adeguati di verifica dell'età (c.d. *age verification*) in grado di garantire l'effettività dei divieti all'uso dei sistemi di intelligenza artificiale da parte dei minori. Il Garante formula alcuni rilievi anche con riferimento alle norme settoriali e, infine, chiede un rafforzamento del proprio ruolo nell'elaborazione della Strategia nazionale per l'intelligenza artificiale.

La Commissione UE, con un documento del 5 novembre 2024[106], invece, ha messo in evidenza le incongruenze tra le disposizioni nazionali contenute nel disegno di legge in esame e l'*AI Act.*

Infatti, la Commissione rileva che, nonostante il testo del disegno di legge affermi esplicitamente la piena conformità della disciplina con la legislazione dell'Unione, vi sono numerosi punti di disallineamento che sono in contrasto con l'esigenza di armonizzare il quadro normativo europeo.

In primo luogo, la Commissione rileva che la proposta normativa italiana si basa su un approccio troppo restrittivo e rigido sull'impiego dell'IA, in particolare nelle professioni intellettuali e nell'attività giudiziaria, e su una tecnica legislativa imprecisa, ambigua e lacunosa, che rischia di sovraccaricare il settore sanitario con obblighi eccessivi.

Inoltre, la Commissione osserva che la designazione dell'AgID e dell'ACN come autorità competenti per l'intelligenza artificiale, non garantisce una sufficiente autonomia e indipendenza dall'Esecutivo e, dunque, non tutela da possibili conflitti di

[105] Garante per la protezione dei dati personali, *Parere su uno schema di disegno di legge recante disposizioni e deleghe in materia di intelligenza artificiale* - 2 agosto 2024 [10043532], https://www.garanteprivacy.it/

[106] COMMISSIONE EUROPEA parere C (2024) 7814

interesse e non assicura che le decisioni siano prese in modo trasparente e obiettivo. La mancanza di indipendenza, quindi, mina la credibilità della supervisione italiana sull'IA e rischia di compromettere l'efficacia della normativa in un settore così delicato.

Il disegno di legge italiano prevede, altresì, che, per esigenze di trasparenza, i contenuti audiovisivi e radiofonici generati o modificati tramite intelligenza artificiale, rechino un segno identificativo "*IA*" per chiarirne la provenienza artificiale. Tale obbligo è previsto anche per contenuti audio e video su piattaforme di streaming. A tal proposito la Commissione Europea ritiene che i citati obblighi sono eccessivi e ridondanti rispetto a quelli già previsti dal Regolamento UE sull'intelligenza artificiale. L'obbligo di visibilità del segno "*IA*", peraltro, potrebbe creare un eccesso di regolamentazione, che complicherebbe l'adeguamento delle piattaforme alle normative europee, generando confusione.

Pertanto, la Commissione suggerisce di inserire esplicitamente nel disegno di legge un riferimento all'*AI Act*, per garantire la coerenza con la normativa europea e di riprendere le medesime definizioni dei diversi modelli di IA, previste dal Regolamento europeo. Inoltre, si suggerisce di chiarire la definizione di "*dati critici*" per i quali è prevista una localizzazione dei data center sul territorio nazionale. Essi, secondo la Commissione, devono essere solo quelli connessi agli interessi di sicurezza nazionale, per evitare che vengano adottate misure che possano apparire sproporzionate rispetto alla necessità di protezione dei dati.

Con riferimento agli obblighi informativi che il disegno di legge prevede per gli operatori di IA in ambito sanitario, la Commissione suggerisce di limitarsi a comunicare ai cittadini che si stanno usando sistemi di AI.

Si chiede, altresì, di eliminare le restrizioni all'uso di sistemi di IA non "ad alto rischio" nelle professioni intellettuali, poiché tali limitazioni potrebbero entrare in conflitto con le disposizioni europee che permettono l'uso dei citati sistemi a condizione che non vi siano rischi significativi per la sicurezza o i diritti fondamentali.

Anche con riferimento all'attività giudiziaria, la Commissione invita ad allineare la norma del disegno di legge con quella del Regolamento europeo, ampliando l'ambito di utilizzo dell'IA.

In merito all'organica definizione della disciplina nei casi di uso di sistemi di intelligenza artificiale per finalità illecite, la Commissione europea ricorda che il

Regolamento prevede specifiche disposizioni in materia di sanzioni per violazioni del regolamento da parte degli operatori. Pertanto, non è necessaria alcuna delega al Governo, sul punto.

In conclusione, sebbene il parere della Commissione UE non sia ostativo, le raccomandazioni formulate richiedono un esame approfondito da parte delle istituzioni italiane. Il Parlamento, infatti, deve valutare come integrare i suggerimenti per evitare conflitti con le normative europee e assicurare che la legge sull'AI rispetti i principi di trasparenza, sicurezza e protezione dei diritti fondamentali[107].

Mentre il disegno di legge è ancora in discussione, l'AgID che, come si è detto, è designata come una delle due Autorità nazionali per l'intelligenza artificiale, sta elaborando linee guida per l'introduzione dell'intelligenza artificiale nella pubblica amministrazione: il piano triennale dell'Agenzia prevede obiettivi ambiziosi, tra cui la definizione di linee guida per favorire l'adozione dell'IA, per bandire gare e appalti dedicati e per sviluppare le prime applicazioni pratiche.

[107] C.MORELLI, *Intelligenza artificiale: la Commissione Ue richiama l'Italia sul Ddl. Occorre rispettare il dettato dell'AI Act senza ulteriori limitazioni*, 4.12.2024, Altalex.it

CAPITOLO II

La giustizia predittiva

Sommario: 2.1 I sistemi predittivi. -2.2 Esperienze e progetti di giustizia predittiva. -2.3 La disciplina della giustizia predittiva: la Carta Etica del 2018, le linee guida dell'UNESCO del 2024 e l'*AI Act*.

2.1 I sistemi predittivi.

La *Carta etica europea sull'utilizzo dell'intelligenza artificiale nei sistemi giudiziari e negli ambiti connessi*[108] adottata dalla Commissione europea per l'efficienza della giustizia del Consiglio d'Europa (CEPEJ) nel 2018 si basa su uno studio approfondito dell'utilizzo dell'intelligenza artificiale nei sistemi giudiziari dal quale si evince che le principali applicazioni riguardano i motori di ricerca giurisprudenziali avanzati, la risoluzione delle controversie online e l'assistenza nella redazione di atti. Inoltre, l'intelligenza artificiale è utilizzata per la categorizzazione dei contratti, per l'individuazione delle clausole contrattuali divergenti o incompatibili e per la creazione di chatbots che informano le parti in lite o le sostengono nel procedimento giudiziario.
L'intelligenza artificiale è, altresì, impiegata per effettuare l'analisi predittiva[109] che è uno dei metodi e delle tecniche di data analytics, adoperate per estrarre valore dai dati, assieme all' analisi descrittiva e a quella prescrittiva. In particolare, l'analisi predittiva va oltre la semplice descrizione dei fatti storici e, attraverso modelli probabilistici, cerca di svelare le relazioni che intercorrono tra dati diversi e di estrapolarle opportunamente, in modo da anticipare uno scenario futuro.

[108] COMMISSIONE EUROPEA PER L'EFFICIENZA DELLA GIUSTIZIA, *Carta etica europea sull'utilizzo dell'intelligenza artificiale nei sistemi giudiziari e negli ambiti connessi*, https://rm.coe.int/carta-etica-europea-sull-utilizzo-dell-intelligenza-artificiale-nei-si/1680993348, p.15.
[109] La data analytics si basa, in primo luogo, sulla descriptive analytics o analisi descrittiva che descrive i dati passati per capire e interpretare una situazione, rispondendo alla domanda generica "*che cosa è successo*" attraverso la predisposizione di tabelle e grafici, arricchiti dall'uso di statistiche riassuntive, semplici trasformazioni e aggregazioni di valori. Dopo aver stabilito cosa è accaduto in passato, i metodi di predictive analytics o analisi predittiva, sono impiegati per rispondere alle domande: "*perché è successo?*" e "*che cosa succederà?*". Infine, la prescriptive analytics o analisi prescrittiva indica direttamente quali azioni realizzare per raggiungere un determinato obiettivo, rispondendo alla domanda: "*che cosa fare?*".

Tale tipologia di analisi è alla base della "*giustizia predittiva*" che, secondo un'autorevole definizione[110], consiste nella «*capacità attribuita alle macchine di mettere in moto rapidamente in un linguaggio naturale il diritto pertinente per trattare un caso, di contestualizzarlo secondo specifiche caratteristiche (del luogo, della personalità dei giudici, degli studi legali, ecc.) e di anticipare la probabilità delle decisioni che potrebbero essere prese*»[111].

La giustizia predittiva si avvale di tecnologie che, attraverso l'esame di norme e precedenti giurisprudenziali, aiutano gli avvocati e le parti a valutare le possibilità di successo di una controversia, prevedendone il risultato finale e sono d'ausilio al magistrato giudicante, suggerendogli la decisione più coerente. La funzione predittiva, dunque, consiste nel mettere in correlazione i dati. Il sistema artificiale riceve i dati, individua delle regolarità tra quelli in input e quelli in output e genera modelli tratti in via induttiva come fonte di conoscenza per situazioni future[112].

La giustizia predittiva si basa sul principio secondo cui il diritto oggettivo può assolvere alla propria funzione di strumento di garanzia della convivenza di un gruppo sociale, solo se è certo e conoscibile e l'interpretazione delle norme è prevedibile e stabile[113].

L'esigenza della certezza del diritto, infatti, ha creato dei punti di contatto tra il pensiero matematico e la conoscenza giuridica. A più riprese, sin dal Medioevo si è cercato di definire strutture semantiche e di formalizzare un calcolo logico applicabile anche a questioni di natura giuridica, in modo da renderle computabili e prevedibili e sottrarle all'alea della soggettività umana[114].

---

[110] A. GARAPON, J. LASSEGUE, *La giustizia digitale. Determinismo tecnologico e libertà*, Il Mulino Bologna, 2021, p.28. Gli autori utilizzano l'espressione "*giustizia predittiva*" come sinonimo di "*giustizia digitale*".

[111] A. GARAPON, J. LASSEGUE, *Ivi*, p.28. Secondo gli Autori si tratta di un fenomeno che comprende strumenti finalizzati a rendere il diritto più prevedibile, che «*non vuole togliere il lavoro agli avvocati ma permettere loro di essere migliori; non intende indebolire la fiducia nella giustizia, ma aumentarla*».

[112] COMMISSIONE EUROPEA PER L'EFFICIENZA DELLA GIUSTIZIA, *Carta etica europea sull'utilizzo dell'intelligenza artificiale nei sistemi giudiziari e negli ambiti connessi*, cit., p. 47 Lo Studio allegato alla carta precisa che «*il termine 'predittivo' utilizzato dalle società di legal tech è tratto dalle branche della scienza (principalmente la statistica) che consentono di predire risultati futuri grazie all'analisi induttiva. Le decisioni giudiziarie sono trattate al fine di scoprire correlazioni tra i dati in ingresso (criteri previsti dalla legge, fatti oggetto della causa, motivazione) e i dati in uscita (decisione formale relativa, per esempio, all'importo del risarcimento). Le correlazioni che sono giudicate pertinenti consentono di creare modelli che, qualora siano utilizzati con nuovi dati in ingresso (nuovi fatti o precisazioni introdotti sotto forma di parametri, quali la durata del rapporto contrattuale), producono secondo i loro sviluppatori una previsione della decisione (per esempio, della forbice risarcitoria)*»

[113] S. MAZZAMUTO, *Manuale di diritto privato*, Giappichelli Torino p.11.

[114] A. TRIGGIANO, *Gli algoritmi predittivi nel processo. Riflessioni storico giuridiche*, ESI, Napoli, 2023, p.53. Secondo l'Autrice, Raimondo Lullo, autore della *Ars Magna*, per la prima volta ha ritenuto

Sulla certezza del diritto, peraltro, si basa anche la giurimetria, disciplina che, come si è detto[115], negli anni '50 del 1900 ha promosso l'utilizzo dei computer in ambito giuridico e, soprattutto, nel processo.

L'evoluzione della ricerca sull'intelligenza artificiale negli anni '80 ha permesso di elaborare sistemi esperti da applicare al diritto, anche per finalità predittive[116]. In particolare, alcuni sistemi esperti si basavano sulle regole e i concetti e altri, sui casi. I primi erano impiegati in contesti in cui andavano applicate numerose regole ben specificate e non controverse, come, ad esempio, quelle di sicurezza e previdenza sociale e quelle tributarie. In tali casi i sistemi esperti potevano garantire l'applicazione diligente delle regole rilevanti per il caso concreto. Essi non erano in grado di fornire un valido supporto laddove i fatti, i concetti o le regole fossero controversi, oppure in ambiti in cui non era possibile codificare in anticipo tutte le situazioni di fatto. I sistemi esperti avevano capacità predittive ed erano in grado di applicare le regole e i concetti di cui erano a conoscenza, al caso concreto, anticipando la decisione futura del decisore competente. Tuttavia, l'affidabilità del risultato dipendeva dall'attività degli esperti che avevano il compito di inserire nel sistema le conoscenze e i relativi aggiornamenti e dalla loro capacità di qualificare i fatti così come avrebbe fatto il decisore.

Come si è detto, vi erano anche sistemi esperti basati sulla rappresentazione di casi. La base di conoscenza di tali sistemi era costituita da un insieme di precedenti descritti, ciascuno, con il relativo esito e attraverso l'individuazione di un insieme di fattori, cioè di circostanze a favore o contro quell'esito. Sottoponendo alla macchina i nuovi casi

---

che il diritto potesse essere considerato come una entità "calcolabile" non solo nelle premesse, ma anche nelle conclusioni giudiziarie che, quindi, sono prevedibili poiché finite. Successivamente, Pietro Ramo, filosofo francese vissuto nel XVI secolo, andò alla ricerca di un nuovo metodo che garantisse una esposizione ragionata della materia scientifica in genere e, del diritto, in particolare. Leibniz, nel XVII secolo, si dedicò alla costruzione del *calculus ratiocinator* per risolvere, non solo i problemi filosofici, ma anche quelli giuridici. Egli riteneva che «*tutte le questioni di diritto puro sono definibili con certezza geometrica*» e, dunque, prevedibili. Successivamente, Nicholaus Bernoulli, Nicolas de Condorcet, Pierre Simon Laplace, Siméon-Denis Poisson elaborarono formule statistico-matematiche per giungere alla previsione della sentenza per ridurre il margine di errore del magistrato chiamato a decidere le sorti dell'imputato. A Max Weber, invece, si deve l'idea che lo sviluppo dell'economia capitalista impone che il diritto sia calcolabile secondo regole razionali. Le strategie di investimento imprenditoriali, infatti, richiedono che il mondo sia prevedibile e, con esso, anche il diritto. M.LUCIANI, *La decisione giudiziaria robotica*,in A. CARLEO(a cura di), *Decisione robotica,* Il Mulino Bologna, 2019, p.76. L'Autore rileva che tra la fine del Settecento e la fine dell'Ottocento si afferma la tendenza alla progressiva spersonalizzazione della figura del giudice le cui prestazioni devono dipendere esclusivamente dalla professionalità, dalla sapienza e dalle qualità intellettuali.

[115] Cfr. *Supra* cap.I

[116] A. SANTOSUOSSO, G. SARTOR, *La giustizia predittiva: una visione realistica,* in *Giurisprudenza Italiana*, Luglio 2022, p.1765.

rappresentati mediante un insieme di fattori, era possibile ottenere indicazioni sulla possibile decisione, attraverso l'analogia. Anche tali sistemi potevano fare previsioni che erano inevitabilmente condizionate dalla capacità degli esperti di descrivere i precedenti cogliendo correttamente i fattori che ne avevano determinato la decisione[117].

I sistemi in esame si sono sviluppati soprattutto negli ordinamenti giuridici dei paesi americani ed anglosassoni di common law dove il precedente giurisprudenziale è vincolante ai fini della decisione della quaestio iuris. Nei paesi di civil law, invece, tali sistemi hanno svolto una funzione di mero supporto all'attività del giudice, degli avvocati e dei cittadini.

L'avvento degli algoritmi di elaborazione del linguaggio naturale (NLP)[118], che permettono a computer e dispositivi digitali di riconoscere, comprendere e generare testo e parlato, e degli algoritmi di apprendimento automatico e di deep learning, ha consentito di costruire sistemi di intelligenza artificiale che, come rileva la citata Carta Etica della CEPEJE, trattano le decisioni giudiziarie. Nella Carta si legge*: «Nella maggior parte dei casi, l'obiettivo di tali sistemi non è la riproduzione di un ragionamento giuridico, bensì l'individuazione delle correlazioni tra i diversi parametri di una decisione (per esempio, in una domanda di divorzio, la durata del matrimonio, il reddito dei coniugi, la presenza di adulterio, l'importo dell'assegno di mantenimento concesso, ecc.) e, mediante l'utilizzo dell'apprendimento automatico, dedurne uno o più modelli. Tali modelli sarebbero successivamente utilizzati per "predire" o "prevedere" una futura decisione giudiziaria*»[119].

I sistemi basati sull'apprendimento supervisionato ricevono da un istruttore umano o prelevano direttamente dal web i casi passati e le soluzioni. Il compito dell'intelligenza artificiale consiste nell'individuare il nesso logico e applicarlo ad altri casi analoghi. L'analisi predittiva, dunque, anticipa le decisioni che di fatto i giudici potrebbero adottare, sulla base di un calcolo probabilistico.

In caso di apprendimento per rinforzo, invece, il sistema apprende dai risultati delle proprie azioni, attraverso ricompense o penalità che sono collegate ai risultati stessi.

---

[117] A. SANTOSUOSSO, G. SARTOR, *Ivi*, p.1766.

[118] I sistemi di Natural Language Processing (NLP) sono in grado di analizzare, rappresentare e comprendere il linguaggio naturale a partire da dati o documenti forniti in input, per tradurlo oppure per produrre testo in modo autonomo.

[119] COMMISSIONE EUROPEA PER L'EFFICIENZA DELLA GIUSTIZIA, *Carta etica europea sull'utilizzo dell'intelligenza artificiale nei sistemi giudiziari e negli ambiti connessi*, https://rm.coe.int/carta-etica-europea-sull-utilizzo-dell-intelligenza-artificiale-nei-si/1680993348, p.23

Pertanto, nell'ambito della giustizia predittiva sistemi di tal genere richiedono la valutazione positiva o negativa delle decisioni assunte dal sistema. Per esempio, la soluzione individuata dal sistema è valutata positivamente se il giudice umano la condivide, mentre è valutata negativamente se è rigettata[120].

I sistemi basati sull'apprendimento non supervisionato, infine, imparano senza ricevere istruzioni e sono utili per riunire insiemi di oggetti che presentano somiglianze o connessioni rilevanti come, ad esempio, i documenti che riguardano gli stessi oggetti o problemi, le persone con caratteristiche simili. Pertanto, tali sistemi sono utili per compiti ausiliari rispetto alla decisione del giudice, ad esempio, nell'ambito di un'indagine, per raggruppare i documenti elettronici disponibili, oppure per raccogliere i casi simili ed individuare le connessioni tra certi reati e l'uso di droghe o di armi [121].

Gli strumenti forniti dalla tecnologia dell'intelligenza artificiale rappresentano certamente una grande opportunità per rendere più efficiente l'apparato giudiziario[122]. Tuttavia, essi portano con sé altrettanti rischi.

Più nello specifico i sistemi di giustizia predittiva potrebbero garantire una maggiore certezza del diritto intesa in termini di calcolabilità e accrescere l'uniformità e la coerenza degli orientamenti giurisprudenziali, riducendo i casi di giudicati contrastanti. Inoltre, le sentenze potrebbero essere sottratte all'errore umano e ai rischi di parzialità, garantendo la neutralità delle decisioni. La possibilità di conoscere preventivamente il probabile esito di un'eventuale azione giudiziaria, poi, consentirebbe di ridurre i tempi della giustizia, rendendo più efficiente il processo decisionale, riducendo il numero delle controversie e dando impulso anche agli strumenti di *alternative dispute resolution* (ADR) come, ad esempio, la mediazione che, peraltro, potrebbe svolgersi da remoto [123].

Tuttavia, l'impiego degli strumenti di giustizia predittiva rischia di ostacolare l'evoluzione della giurisprudenza e del diritto[124]. Come si è detto, gli algoritmi di *machine learning* e di *deep learning* si basano su una logica di stampo prettamente

[120] A. SANTOSUOSSO, G. SARTOR, *La giustizia predittiva: una visione realistica*, cit., p.1767.
[121] A. SANTOSUOSSO, G. SARTOR, *Ivi*, p.1768.
[122] M. LIBERTINI, M.R. MAUGERI, E. VINCENTI, *Intelligenza artificiale e giurisdizione ordinaria. Una ricognizione delle esperienze in corso*, in Astrid Rassegna, 16/2021, p. 4; A. BALSAMO, *L'impatto dell'intelligenza artificiale nel settore della giustizia*, in *Sistema penale*, 2024, p.6.
[123] M.R. COVELLI, *Dall'informatizzazione della giustizia alla decisione robotica? Il Giudice del merito*, in A. CARLEO (a cura di), *La decisione robotica*, Il Mulino, Bologna, 2019, p.126. La prevedibilità delle decisioni giudiziarie è un valore che ha le sue radici nella Costituzione (artt.3, 24,97, 101 Cost) , genera sicurezza ed affidamento per l'individuo, riduce il contenzioso e si ripercuote positivamente sull'efficacia degli strumenti deflattivi.
[124] A. BALSAMO, *L'impatto dell'intelligenza artificiale nel settore della giustizia*, cit., p.7.

statistico-probabilistico. Pertanto, essi individuano una correlazione osservata nella maggior parte dei casi su cui sono stati addestrati e non comprendono i meccanismi causali del fenomeno. Gli algoritmi, dunque, formulano predizioni sulla base di quanto già accaduto in passato e non lasciano spazio per l'ingresso di nuovi orientamenti giurisprudenziali e, attraverso essi, per l'adeguamento del diritto alle esigenze sociali del momento.

Inoltre, gli strumenti in esame consentirebbero un'impropria profilazione dei giudici che avrebbero difficoltà a distaccarsi dai propri orientamenti a fronte della percezione pubblica di una serie di decisioni prese in precedenza ed espressive di un indirizzo che rischia di produrre una vera e propria prigionia del passato.

Occorre, altresì, considerare che non esistono macchine infallibili e anche gli algoritmi, in quanto prodotti dagli esseri umani, come si è detto, possono sbagliare. Le macchine o i computer, infatti, non sono del tutto imparziali. I sistemi di intelligenza artificiale si basano su algoritmi che possono essere condizionati da "*bias*", da pregiudizi, mutuati dai *dataset* di addestramento[125].

Le potenziali mancanze relative ai dati usati dall'intelligenza artificiale possono compromettere la qualità, l'efficacia e l'efficienza dei risultati e rischiano anche di incrementare il numero di decisioni e azioni discriminatorie nei confronti di categorie e gruppi di persone già emarginate all'interno della società. Pertanto, l'apparente neutralità dei sistemi di intelligenza artificiale rischia di celare pratiche discriminatorie che potrebbero essere causate da un dataset viziato e pregiudizievole, rendendo ancora più complessa l'individuazione, la contestazione e la sanzione delle discriminazioni perpetrate attraverso l'uso di tale tecnologia[126].

Un ulteriore problema connesso all'utilizzo degli strumenti predittivi è costituito dal fenomeno della *black box* che li rende poco trasparenti. Gli algoritmi di *machine learning* e di *deep learning*, infatti, non consentono ai destinatari di una decisione che sia anche solo parzialmente automatizzata di conoscere almeno la logica di fondo che l'ha ispirata. Pertanto, è impossibile l'esercizio del controllo e della verifica dei motivi alla base della decisione e della correttezza del processo eseguito dall'algoritmo. Ogni decisione assunta con l'impiego dell'intelligenza artificiale, quindi, sarebbe viziata in

[125] M.LUCIANI, *La decisione giudiziaria robotica,* in A. CARLEO (a cura di), *La decisione robotica*, Il Mulino, Bologna, 2019, p.79.
[126] A. BALSAMO, *L'impatto dell'intelligenza artificiale nel settore della giustizia*, cit., p.7.

punto di motivazione, con il rischio di essere automaticamente illegittima. Inoltre, l'uso dell'intelligenza artificiale a fini decisionali potrebbe determinare una deresponsabilizzazione, non solo del difensore, ma anche del giudice che potrebbe tendere a conformarsi acriticamente ai risultati forniti dall'algoritmo. Ciò causerebbe, altresì, la perdita della dimensione emotiva e valoriale del giudizio[127].

Infine, non va trascurato il rischio connesso alla natura prevalentemente privata dei soggetti che, negli ultimi anni, hanno finanziato e sostenuto lo sviluppo e la realizzazione dei sistemi di intelligenza artificiale. A tal proposito va rilevato che i sistemi in esame sono dei prodotti tutelati da diritti di proprietà intellettuale, commercializzabili su mercati concorrenziali. I produttori hanno un interesse economico ad acquisire e conservare una posizione di vantaggio e di prevalenza nel mercato globale e, quindi, tengono segreti i meccanismi di creazione e di funzionamento dei vari sistemi che non possono essere sottoposti a verifiche. Pertanto, in primo luogo, ciò contribuisce a rendere opache le decisioni assunte con l'ausilio dei sistemi in esame. Inoltre, vi è il rischio che lo sviluppo delle tecnologie di intelligenza artificiale non sia inclusivo, escludendo coloro che non sono in grado di sostenerne economicamente gli elevati costi. Infine, come si è detto, i produttori privati, attraverso l'analisi dei Big data e l'elaborazione di modelli predittivi e decisionali, possono conoscere i comportamenti e i desideri delle persone, influenzandone e indirizzandone le decisioni e le azioni. Si rischia, dunque, che pochi soggetti privati, tutelati dai diritti di proprietà intellettuale, detengano il potere di sorvegliare e di manipolare le decisioni legate alla sfera privata e alla dimensione pubblica di una collettività inconsapevole, compromettendo la tutela degli interessi, dei diritti e delle libertà delle persone[128].

2.2 Esperienze e progetti di giustizia predittiva.

I sistemi di giustizia predittiva basati sull'apprendimento automatico e sul riconoscimento del linguaggio naturale presentano diverse combinazioni di una molteplicità di funzioni. Essi, infatti, possono prevedere il risultato finale di un caso giudiziario oppure eventi o circostanze che possono contribuire a determinarlo. Inoltre, possono anticipare il comportamento del giudice oppure offrirgli indicazioni di comportamento.

[127] A. BALSAMO, *Ivi* p.8.
[128] A. SANTOSUOSSO, G. SARTOR, *La giustizia predittiva: una visione realistica*, cit., p.1781.

La previsione si può basare su elementi normativamente rilevanti per il caso specifico, che il giudice dovrebbe considerare per adottare o giustificare la propria decisione oppure su elementi normativamente irrilevanti, che il decisore non dovrebbe prendere in considerazione. Inoltre, può essere basata sul contenuto testuale della sentenza o di un'altra decisione del giudice, oppure su dati non testuali. Infine, il sistema può essere in grado di fornire una spiegazione delle sue decisioni oppure può rimanere opaco e non riuscire a farlo[129].

Attualmente in campo giudiziario sono adottate solo tecniche di intelligenza artificiale capaci di apprendere da grandi volumi di dati in maniera efficiente e con i maggiori livelli teorici di precisione predittiva. Le prime applicazioni concrete di algoritmi di intelligenza artificiale in ambito giudiziario sono state sperimentate in Cina[130] dove, nella Procura di Shanghai Pudong, è stato implementato il primo giudice robot che, secondo i ricercatori della Chinese Academy of Science, è in grado di prendere decisioni sulla colpevolezza con un'accuratezza del 97%. Inoltre, è stata creata la Beijing Internet Court, una Corte online che, attraverso un centinaio di robot, fornisce supporto ai giudici nel processo decisionale, svolgendo compiti ripetitivi come la ricezione dei ricorsi e l'analisi giurisprudenziale di casi simili. Il sito della Corte rende liberamente accessibili quasi tutti i documenti delle sentenze pronunciate dai tribunali dal 2013[131].

In Canada, in ambito civile, è impiegato un sistema di risoluzione delle controversie online di basso valore.

In Olanda sono utilizzati software per l'elaborazione del rischio di recidiva dell'indagato e si sta cercando di automatizzare alcuni processi decisionali. Inoltre, è stato elaborata dall'Università di Twente e dallo Hague Institute for Internationalisation of the Law una piattaforma per la gestione on line dei casi di mediazione e di soluzione extragiudiziale delle controversie di carattere civile[132].

---

[129] A. SANTOSUOSSO, G. SARTOR, *Ibidem.*

[130] M. GUO, *Internet Court's challenges and future in China*, in *Computer Law & Security Review*, 40, aprile 2021. In Cina sono liberamente accessibili quasi tutti i documenti delle sentenze pronunciate dai tribunali dal 2013.

[131] A. BALSAMO, *L'impatto dell'intelligenza artificiale nel settore della giustizia*, cit., p.8.

[132] E.E. ARTESE, E.V. CONFORTI, D. GHIDELL, *La giustizia predittiva: potenzialità e casi di studio, in* AA.VV., *Atti del Convegno "Nuovi scenari della giustizIA",* (a cura di Maria Novella Campagnoli e Massimo Farina), Ottobre 2022, in *Rivista elettronica di Diritto, Economia, Management*, 2024, 1, p.30.

In Francia, nelle Corti d'appello di Rennes e Douai, a seguito della dematerializzazione dei fascicoli di ultima istanza, è stata sperimentata la piattaforma Predictice. Tale piattaforma, creata da una start up privata, è nata per aiutare gli avvocati nell'elaborazione delle migliori strategie processuali, esplicitando la probabilità di accoglimento o rigetto di una determinata argomentazione.

In Gran Bretagna, nel 2016, la University College di Londra ha creato un sistema di intelligenza artificiale in grado di prevedere, con un grado di affidabilità del 70%, le decisioni della Corte Europea dei Diritti Umani. L'Università di Edimburgo ha elaborato un sistema denominato ADVOKATE che consente la valutazione della competenza ed affidabilità dei testimoni.

In Austria, l'intelligenza artificiale è utilizzata nei tribunali per effettuare la lettura rapida dei documenti, per la classificazione e l'attribuzione degli atti e per monitorare l'attività degli uffici giudiziari.

In Estonia è stato adottato un software per il calcolo del rischio della recidiva e il Ministero della Giustizia, alla fine del 2019, ha avviato un progetto pilota per l'impiego dell'intelligenza artificiale per risolvere controversie di valore fino a 7.000 euro. Le decisioni robotiche sono appellabili dinanzi al giudice-persona[133].

Anche negli Stati Uniti, da tempo sono utilizzati software predittivi del rischio di recidiva, in primo luogo, nella fase preliminare del giudizio per la determinazione della cauzione. Inoltre, tali sistemi sono applicati in fase pre-decisoria per valutare la possibilità di definire il procedimento con una sentenza simile a quella di messa alla prova. Infine, essi sono impiegati nella fase esecutiva per la valutazione della concessione della libertà condizionale.

L'impiego di algoritmi predittivi nei percorsi decisionali giudiziali negli Stati Uniti ha suscitato un acceso dibattito nel 2016 con il caso Loomis vs. Wisconsin[134] che ha messo in luce le problematiche giuridiche e i rischi per i diritti fondamentali, connessi all'utilizzo di tali strumenti in fase decisoria.

Il caso ha coinvolto Eric Loomis, arrestato nel 2013 per essere stato alla guida di un'auto coinvolta in una sparatoria. Durante il processo gli sono stati addebitati cinque capi d'accusa: messa in pericolo della sicurezza, tentativo di fuga da un ufficiale del traffico, guida di un veicolo senza il consenso del proprietario, possesso di un'arma da

[133] E.E. ARTESE, E.V. CONFORTI, D. GHIDELL, *Ivi*, cit., p.31.
[134] State of Wisconsin v. Eric L. Loomis, 881 N.W.2d 749 (Wis. 2016).

fuoco da parte di un pregiudicato e possesso di un fucile a canna corta o di una pistola.[135]

L'imputato ha deciso di patteggiare, dichiarandosi colpevole solo dei due reati minori e cioè del tentativo di fuga da un agente della polizia stradale e dell'uso di un veicolo senza il consenso del proprietario. La Corte del Wisconsin ha accettato l'ammissione di colpevolezza di Loomis e ha ordinato un Presentence Investigation Report (PSI) che consiste in una relazione dei risultati delle investigazioni condotte sulla storia personale dell'imputato, prima del patteggiamento, per verificare circostanze utili a modulare la severità della pena. La relazione era redatta anche sulla base dei risultati elaborati dal software COMPAS (Correctional Offender Management Profiling for Alternative Sanctions), uno strumento prodotto da una società statunitense per calcolare il rischio di recidiva e la pericolosità sociale in base a dati statistici, ai precedenti giudiziari, alle risposte a un questionario fornite dall'imputato e a una serie di variabili non note, né conoscibili in quanto coperte da proprietà intellettuale. La Corte che, per la prima volta, ha usato il COMPAS in fase decisoria ha condannato l'imputato Loomis a sei anni di reclusione e cinque anni di supervisione, considerando il casellario penale e i risultati elaborati dal citato sistema di intelligenza artificiale.

Loomis ha presentato un'istanza di revisione della pena sostenendo che l'utilizzo di un algoritmo predittivo da parte del giudice avesse violato le garanzie del giusto processo. In particolare, Loomis sosteneva che il software COMPAS rendesse le valutazioni del rischio finali, opache e non verificabili scientificamente. Tuttavia, il Tribunale circondariale ha rigettato l'istanza di revisione e ha confermato la condanna, rilevando che la storia pregressa di Loomis sarebbe stata valutata allo stesso modo, anche senza l'utilizzo di COMPAS.

Loomis ha impugnato la decisione innanzi alla Corte d'Appello che ha rimesso la questione alla Corte Suprema del Wisconsin. La Corte ha dichiarato la legittimità dell'utilizzo di algoritmi per la valutazione del rischio di recidiva anche in fase di cognizione precisando, tuttavia, che lo strumento non può essere l'unico elemento su

[135] F. LAGIOIA, R. ROVATTI, G. SARTOR, *Fairness through group parities? the case of COMPAS-SAPMOC*, in *AI and Society*, 2022; M. SCIACCA, *Algoritmo e giustizia alla ricerca di una mite predittività*, in *Persona e Mercato*, 2023, 1, p.10; C. SALAZAR, *Judex ex machina? Note su giustizia, giudici e intelligenza artificiale*, in *Consulta on line* 2021, p.924.

cui si fonda una pronuncia di condanna, ma solo un supplemento al processo decisionale giudiziario. La Corte ha posto numerosi limiti all'utilizzo del COMPAS stabilendo che esso non potesse essere impiegato per determinare l'opportunità o meno di comminare una pena detentiva o per calcolarne la durata. Inoltre, l'uso del citato strumento doveva essere accompagnato da una motivazione indipendente e tutti i rapporti generati dal sistema, dovevano contenere un avvertimento sull'utilità limitata dell'algoritmo.

La Corte, inoltre, ha evidenziato il rischio che COMPAS attribuisse importanza sproporzionata ad alcuni fattori come il contesto familiare, il livello di educazione e l'appartenenza a un gruppo etnico. Esso era un sistema opaco in quanto, il produttore, per conservare il proprio vantaggio competitivo, non consentiva di controllare il codice sorgente e la base dati di addestramento e di verificare se l'algoritmo fosse distorto oppure no. Peraltro, alcuni studiosi americani[136] avevano analizzato il funzionamento di COMPAS con uno studio dal quale emergeva che l'algoritmo soffriva dei pregiudizi razziali allo stesso modo dell'essere umano.

Secondo la Corte Suprema del Wisconsin, per garantire l'accuratezza della valutazione, dunque, il software doveva essere costantemente monitorato e aggiornato sulla base dei cambiamenti sociali. Inoltre, doveva essere utilizzato con accortezza e seguendo specifiche cautele. I tribunali dovevano considerare attentamente i risultati del software con riguardo a ogni specifico individuo, tenendo conto anche di tutti gli altri fattori a disposizione. Al giudice era affidato il compito di valutare se vi fossero possibili profili di illiceità applicativa del software e se tali violazioni potessero ledere il diritto dell'imputato a ricevere un giusto processo.

Anche Italia sono stati applicati degli algoritmi predittivi per integrare l'attività investigativa umana nelle operazioni di polizia preventiva. Al riguardo, meritano di essere menzionati i sistemi KeyCrime, sviluppato dalla Polizia di Milano e utilizzato per prevedere le rapine nell'area metropolitana, e X-LAW, sviluppato dalla Polizia di

[136]J. LARSON et al., *How We Analyzed the COMPAS Recidivism Algorithm*, in *ProPublica*, 2016, https://www.propublica.org/article/how-we-analyzed-the-compas-recidivism-algorithm. L'inchiesta è stata realizzata da ProPublica, un'organizzazione non governativa con sede a Manhattan che, mediante il giornalismo investigativo, mira a denunciare abusi di potere da parte di istituzioni pubbliche e private e ha interessato un campione di oltre settemila individui arrestati tra il 2013 e il 2014 nella contea di Broward e i relativi punteggi di rischio calcolati dal COMPAS.

Napoli e applicato dalle forze dell'ordine in diverse regioni italiane per prevedere furti e rapine[137].

KeyCrime è un software di analisi del crimine finalizzato a coadiuvare i dipartimenti di polizia basato su un algoritmo di machine learning che si ispira all'approccio analitico utilizzato dagli investigatori. Si tratta di uno strumento che permette di identificare eventi criminali gravi che potrebbero far parte di una serie, evidenziandone i collegamenti. Esso è compatibile con il sistema giudiziario penale italiano e può essere utilizzato solo per condotte seriali.

Il progetto X-LAW, invece, è un sistema basato su un modello di apprendimento supervisionato che è stato sviluppato dopo un lungo studio criminologico. L'analisi dei dati ha rivelato che tra i tanti crimini che un individuo può commettere, alcuni, tra cui furti, scippi, rapine, borseggi e truffe, sono prevedibili. Infatti, i crimini predatori urbani sono commessi da individui che cercano un profitto immediato e seguono strategie simili in tutti i centri urbani. Si è, altresì, rilevato che i citati crimini avvengono in luoghi che sono caratterizzati dalla presenza di prede e obiettivi allettanti, dove vi sono vie di fuga, rifugi e copertura criminale. Sovrapponendo i crimini consumati e risultanti dalle denunce dei cittadini o dalle informazioni della polizia o delle attività di prossimità alle dinamiche socioeconomiche, si è in grado di decodificare le sequenze spazio-temporali e quindi prevedere i singoli delitti[138]. Il sistema, dunque, fornisce una mappa di rischio che raffigura, ogni trenta minuti, i luoghi e gli orari precisi in cui si potrà consumare un crimine, descrivendo il tipo di crimine, il modus operandi dell'autore, il tipo di preda e di target.

Da qualche anno le Corti d'Appello, i Tribunali e le Università hanno avviato la sperimentazione di diversi progetti pubblici che prevedono l'impiego più o meno preponderante dell'intelligenza artificiale, anche a fini organizzativi. Essi hanno carattere settoriale e non ambiscono ad automatizzare l'intero processo decisionale.

Il progetto "Giustizia predittiva" della Corte d'Appello di Brescia, ad esempio, ha avuto inizio nel 2018 e consiste nella creazione di una piattaforma web in cui sono state raggruppate per aree tematiche alcune decisioni del Tribunale Ordinario di Brescia e

[137] F. CORONA, *La decisione del giudice tra precedente giudiziale e predizione artificiale*, in *Democrazia e diritti sociali*, 2023, 1, p.91; F. DONATI, *Intelligenza artificiale e giustizia*, in *Rivista di associazione italiana costituzionalisti*, 2020, 1, p.418.

[138] F. CORONA, *Ivi*, p.94.

della Corte di Appello di Brescia in materia di diritto societario e industriale, contratti bancari, licenziamenti, diritto contributivo, infortunistica sul lavoro e appalti. La pagina relativa a ciascuna area tematica consente di visualizzare e conoscere come, in precedenza, sono state risolte specifiche controversie e quanto è durato il giudizio[139].

Il progetto di ricerca dell'Università Ca' Foscari Venezia, in collaborazione con la Corte di Appello di Venezia, Deloitte e Unioncamere del Veneto, invece, si basa sulla realizzazione di una piattaforma di intelligenza artificiale che contiene una raccolta giurisprudenziale delle sentenze del distretto veneto, volta a far conoscere agli operatori del diritto e alle imprese i propri orientamenti giurisprudenziali[140]. Un gruppo di ricercatori ha raccolto, esaminato e massimizzato le sentenze relative a controversie di diritto del lavoro, estrapolando da ciascuna di esse una doppia descrizione relativa al fatto e agli enunciati di diritto. In tal modo è stato creato un database ricco e dettagliato, che costituisce il cuore del sistema di giustizia predittiva.

Il dipartimento di intelligenza artificiale della Deloitte ha trasformato questa vasta mole di dati in uno strumento predittivo funzionale. Il processo si articola in due fasi preparatorie fondamentali. La prima fase consiste nell'immissione di un numero notevole di sentenze nel sistema e nella creazione di una mappa cognitiva che permette di categorizzarle e validare i documenti. Nella seconda fase si un esperto legale analizza e valida il materiale per la creazione del database all'interno del tool, che costituisce la base di conoscenza su cui opera l'algoritmo.

Qualsiasi operatore può effettuare una ricerca utilizzando un linguaggio naturale e ottenere un indicatore sintetico che rappresenta l'orientamento prevalente della giurisprudenza rispetto alla tematica specifica richiesta.

Il progetto "Predictive Jurisprudence" realizzato dalla Corte d'Appello di Genova, il Tribunale di Pisa e la Scuola Superiore Sant'Anna di Pisa si caratterizza, invece, per la possibilità di estrarre modelli di ragionamento che soddisfano l'esigenza di spiegabilità e trasparenza delle decisioni dell'algoritmo[141]. Algoritmi di analisi del linguaggio naturale e di machine learning sono utilizzati per realizzare un sistema di lettura e

[139]La documentazione del progetto è rinvenibile sul sito https://giustiziapredittiva.unibs.it/.

[140] La documentazione del progetto è rinvenibile sul sito https://www.corteappello.venezia.it/index.php/giurisprudenza-predittivaper_198.html con una partizione degli orientamenti giurisprudenziali relativi alle aree del diritto di impresa e industriale, diritto del lavoro, diritto in materia bancaria e diritto societario.

[141] La documentazione del progetto è rinvenibile sul sito https://www.predictivejurisprudence.eu/

classificazione automatica delle sentenze estratte dalla banca dati PCT dei Tribunali di Pisa e Genova grazie all'addestramento di un algoritmo. I ricercatori intendono addestrare l'algoritmo in modo che sia in grado di proporre casi simili a quello di volta in volta proposto, decisioni simili, prove simili, ragionamenti simili, regole giuridiche applicate.

Anche la Terza sezione civile della Corte d'Appello di Bari, ha avviato un progetto per rendere accessibili attraverso il sito web, schede tematiche sulla sua giurisprudenza consolidata su materie e casistica ricorrenti, per fornire agli utenti indicazioni circa il prevedibile esito di una possibile controversia in tali materie, nonché sul tempo necessario per giungere alla sentenza[142].

Ai progetti finora elencati si aggiunge quello dell'Università di Bologna ER4JUSTICE – AI4JUSTICE che studia l'utilizzo dell'intelligenza artificiale nei sistemi giudiziari, applicata anche agli strumenti di giustizia predittiva e alla prevenzione della criminalità e del rischio per le persone e gli spazi.[143]

La Corte d'Appello di Milano[144], invece, ha introdotto alcuni sistemi informatici che svolgono un ruolo rilevante e crescente nell'organizzazione dell'ufficio, poiché permettono di verificare con tempestività il raggiungimento degli obiettivi del Piano Nazionale di Resistenza e Resilienza (PNRR), i risultati in divenire delle variabili chiavi del Programma di Gestione e l'andamento dei target tendenziali, nonché di individuare gli obiettivi strutturali cui tendere e corrispondenti ad un buon livello di efficienza. A tal fine sono stati introdotti diversi strumenti informatici.

In primo luogo, è impiegato il simulatore, un sistema che richiama, come modalità di funzionamento, l'intelligenza artificiale.

Tale strumento integra e mette in relazione le diverse fonti statistiche già a disposizione della Corte rendendo disponibile alle sezioni e alla Presidenza un insieme consistente di informazioni sugli andamenti rilevati, sulla produttività e sui principali indicatori di monitoraggio. Con una solida base informativa fattuale, si vuole dare alle sezioni del Tribunale la possibilità di ottimizzare la qualità, i risultati prodotti e il benessere dei magistrati in funzione delle risorse esistenti.

---

[142] La documentazione del progetto è rinvenibile sul sito www.giustizia.bari.it/buone_prassi_4.aspx.

[143] La documentazione del progetto è rinvenibile sul sito https://er4justice.fondazionecrui.it/

[144] G. ONDEI, *L'intelligenza artificiale: la rivoluzione culturale del millennio*, 11ottobre 2024, in rivistaildirittovivente.it

Il simulatore permette di calcolare per l'anno in corso i risultati che l'ufficio potrebbe raggiungere in tema di provvedimenti emessi, oltre ad essere in grado di prevedere diversi scenari sulla base di alcune variabili chiave.

Pertanto, in maniera rapida ed automatica, è possibile monitorare l'andamento dell'attività giurisdizionale e verificarne l'adeguatezza rispetto agli obiettivi, ma anche controllare che venga rispettato il carico esigibile medio di sezione.

Presso la Corte d'Appello di Milano è stato implementato anche un sistema di monitoraggio degli addetti all'Ufficio del processo (A.U.P.), uno strumento informatico alimentato con i dati raccolti attraverso i report mensili e giornalieri, in caso di smart working, relativi al lavoro svolto.

I dati confluiscono su un sistema di sintesi (basato excel) che, a sua volta, costituisce una vera e propria banca dati.

Nel 2021 anche la Corte di cassazione, in collaborazione con la IUSS di Pavia, ha sottoscritto un piano quinquennale per la realizzazione di un incubatore che racchiuda in sé tutta la conoscenza giurisprudenziale e legislativa, per creare uno strumento di facile consultazione per gli operatori del diritto e gli studiosi. L'interessante progetto di ricerca prevede, peraltro, l'utilizzo non solo dell'intelligenza artificiale, ma anche di strumenti di legal analytics. Tali strumenti permetteranno di estrarre e rappresentare la conoscenza giuridica, rinvenire correlazioni implicite e individuare tendenze circa gli orientamenti giurisprudenziali e/o legislativi[145].

In linea generale, dunque, le ricerche sperimentali in corso si basano su una attività di massimazione delle sentenze con livelli di automazione diversificati. I progetti pubblicano il dataset di sentenze massimate su cui si intende costruire, in una fase successiva, un motore di ricerca più o meno avanzato. La base dati giurisprudenziali, teoricamente, serve a produrre effetti deflattivi del contenzioso oppure a indicare, con l'ausilio della statistica avanzata, le tendenze giurisprudenziali.

Dopo l'introduzione del processo civile telematico, il Ministero della Giustizia, nell'ambito dell'attuazione del PNRR, ha promosso dei progetti di ricerca applicata, con lo scopo di estrarre la conoscenza contenuta nel patrimonio documentale e realizzare l'anonimizzazione delle sentenze, l'automazione nell' individuazione del rapporto vittima-autore nei provvedimenti giurisdizionali, la gestione e l'analisi della conoscenza

[145] La documentazione del progetto è rinvenibile sul sito https://www.iusspavia.it/

del processo, un sistema di controllo di gestione del processo e la rilevazione statistica avanzata sui procedimenti civili e penali[146].

Alle Università è stato chiesto di definire i moduli operativi per la costituzione e l'implementazione dell'Ufficio per il Processo[147], di individuare i modelli per la gestione dei flussi in ingresso e degli arretrati presso gli Uffici Giudiziari, di attivare e sperimentare i modelli e i piani relativi alle azioni precedenti e di ridefinire i modelli formativi e consolidare i rapporti tra gli stakeholders. In tale contesto l'intelligenza artificiale rappresenta un prezioso strumento di ausilio per gli addetti all'Ufficio per il Processo. Ad esempio, gli algoritmi di intelligenza artificiale possono essere impiegati per analizzare grandi quantità di dati giurisprudenziali, accelerando la preparazione dei fascicoli e la stesura delle bozze di provvedimenti. Inoltre, l'intelligenza artificiale può aiutare nel monitoraggio delle performance e nel suggerire miglioramenti organizzativi, contribuendo, così, a rendere l'attività giudiziaria più rapida ed efficiente.

Il Ministero della Giustizia, nella circolare del 3 novembre 2021, ha sottolineato l'importanza di istituire servizi trasversali dedicati all'innovazione e alla digitalizzazione

degli uffici giudiziari.

Gli investimenti del PNRR, quindi, hanno consentito di creare la banca dati di merito della giurisprudenza civile che è accessibile sia per i magistrati, che per il pubblico. La parte riservata ai magistrati include milioni di sentenze civili di I e II grado pubblicate a partire dal 2016. Sono escluse le sentenze che riguardano i rapporti di famiglia, i minori e lo stato delle persone in quanto i dati in esse contenuti non sono oscurati. Tale banca dati è integrata con funzioni di intelligenza artificiale che generano sintesi dei provvedimenti. Inoltre, esse sono assistite da una chatbot che risponde alle domande degli utenti in linguaggio naturale. Nella banca dati è, altresì, integrato un sistema di ricerca smart. Essa è aperta anche al pubblico, ma con funzionalità molto limitate[148].

---

[146] M. SCIACCA, *Algoritmo e giustizia alla ricerca di una mite predittività*, cit., p.14.

[147] L'Ufficio per il Processo è una struttura organizzativa prevista dall' art. 16-octies del D.L. n. 179/2012 istituita presso i tribunali ordinari e presso le Corti d'appello con l'obiettivo di garantire la ragionevole durata del processo, attraverso l'innovazione dei modelli organizzativi ed assicurando un più efficiente impiego delle tecnologie dell'informazione e della comunicazione.

[148] La banca dati utilizza la tecnologia Microsoft Azure Open AI per i servizi di intelligenza artificiale LLM e una tecnologia proprietaria, Iride AI, per l'anonimizzazione e pseudonomizzazione dei provvedimenti di area civile. Vi è poi la piattaforma Elasticsearch di Elastic per l'indicizzazione dei testi e dei metadati associati ai provvedimenti inclusi gli abstract e la ricerca full text ei contenuti dei documenti.

Come rileva il Consiglio Superiore della Magistratura (CSM) nella Relazione sulla giustizia telematica approvata il 24 luglio 2024[149] le banche dati integrate con l'intelligenza artificiale non servono solo ad estrapolare e gestire risultati eventualmente utili, ma sono veri e propri estrattori di conoscenza, di supporto ai magistrati nella fase decisionale. Tuttavia, anche il CSM evidenzia i rischi connessi all'impiego di tali strumenti. Pertanto, riservandosi di validare o meno i documenti organizzativi, ricognitivi degli elementi fondanti della governance degli uffici giudiziari, l'organo di autogoverno dei giudici ritiene che sia necessario garantire maggiore trasparenza riguardo ai dati e agli algoritmi utilizzati nei processi decisionali automatizzati e chiede che i fornitori di intelligenza artificiale rendano noti i dati utilizzati durante le fasi di pre-addestramento e addestramento. Il CSM sottolinea, altresì, la necessità di una regolamentazione giuridica nazionale specifica per l'uso dell'intelligenza artificiale in ambito giudiziario che tenga conto della cornice normativa delegata dal Regolamento europeo.

La Corte di cassazione e il Consiglio Nazionale Forense, invece, hanno chiesto e ottenuto dal Ministero della Giustizia l'istituzione dell'Osservatorio permanente per l'uso dell'intelligenza artificiale avvenuta con un decreto ministeriale del 10 luglio 2024.

L'Osservatorio nasce come un luogo di riflessione e approfondimento che coinvolge gli attori fondamentali della giurisdizione e del processo, in cui affrontare tutti i temi che riguardano il rapporto tra intelligenza artificiale e giurisdizione, a partire dalla qualità e sicurezza delle banche dati giuridiche, agli strumenti di supporto dell'attività giurisdizionale e delle professioni[150].

2.3 La disciplina della giustizia predittiva: la Carta Etica del 2018, le linee guida dell'UNESCO del 2024 e l'*AI Act*.

I vantaggi connessi all'utilizzo degli strumenti di intelligenza artificiale in ambito giudiziario hanno fatto sì che essi si diffondessero rapidamente. Da un sondaggio condotto dall'UNESCO nel 2023 emerge che il 93% degli operatori giudiziari ha familiarità con le tecnologie di intelligenza artificiale, e il 44% già le utilizza per attività legate al lavoro. Tuttavia, solo il 9% degli operatori giudiziari intervistati ha riferito che

[149] CSM, Relazione sullo stato della giustizia telematica – 2024, www.csm.it
[150] F. CORONA, *La decisione del giudice tra precedente giudiziale e predizione artificiale*, cit., p.94.

le proprie organizzazioni hanno emanato linee guida o fornito formazione relativa all'intelligenza artificiale.[151] Pertanto, emerge la necessità di stabilire delle regole per evitare che l'utilizzo degli strumenti in esame si riveli lesivo dei diritti fondamentali degli individui.

Come si è detto, la regolamentazione dell'utilizzo dell'intelligenza artificiale, fino alla recente approvazione dell'*AI Act*, è stata essenzialmente demandata a strumenti di soft law. In prima battuta i sistemi di intelligenza artificiale sono stati studiati e regolamentati soprattutto dal punto di vista etico e morale, per l'impatto che il loro uso può avere sulla dignità degli individui titolari dei dati impiegati per addestrarli. Trattare i dati in modo consapevole impone di considerare gli aspetti etici, che si manifestano soprattutto quando essi sono utilizzati in processi decisionali critici come, ad esempio, la valutazione del personale, l'ammissione all'università o le condanne penali.

Tali riflessioni sono alla base del lavoro svolto dalla CEPEJ, un organismo composto da tecnici, rappresentativo dei Paesi membri del Consiglio d'Europa, istituito per testare e monitorare l'efficienza ed il funzionamento dei sistemi giudiziari europei.

Nel 2016 la CEPEJ ha elaborato uno studio approfondito sull'uso dell'intelligenza artificiale nei tribunali europei[152]. Sulla base dei risultati dello Studio, la Commissione ha elaborato le Linee direttrici sulla "cybergiustizia" e, quindi, la *Carta Etica Europea sull'uso dell'intelligenza artificiale nei sistemi giudiziari e in ambiti connessi*, il primo testo europeo non vincolante sui princìpi etici relativi all'uso dell'intelligenza artificiale nell'amministrazione della giustizia.

La Carta etica è stata realizzata dal Gruppo di lavoro sulla qualità della giustizia, composto da sei membri (magistrati, professori universitari, funzionari ministeriali) nominati su proposta dei rispettivi Governi, ed è stata formalmente adottata dalla Commissione durante la sessione plenaria del 3-4 dicembre 2018.

Considerato che l'uso dell'intelligenza artificiale può accrescere l'efficienza e la qualità dei servizi giudiziari e dovrebbe essere incoraggiato, come rileva la Carta, occorre garantire che esso avvenga in modo responsabile, nel rispetto dei diritti fondamentali

[151] Il sondaggio è stato svolto nell'ambito delle attività del programma UNESCO Global Judges Initiative, in collaborazione con esperti internazionali, https://www.unesco.org/en/artificial-intelligence/rule-law

[152] CEPEJ, Studio n. 24, *Rapport thématique: l'utilisation des technologies de l'information par les tribunaux en Europe*, 2016 (dati del 2014).

della persona, enunciati nella Convenzione europea sui diritti dell'uomo e nella Convenzione per la protezione dei dati di carattere personale.

Pertanto, la Commissione stabilisce i principi fondamentali che devono orientare la definizione delle politiche pubbliche relative all'uso dell'intelligenza artificiale nel campo della giustizia[153].

La Carta è destinata agli attori pubblici e privati incaricati di creare e lanciare strumenti e servizi di intelligenza artificiale relativi al trattamento di decisioni e dati giudiziari, nonché ai responsabili di decisioni pubbliche competenti in materia di quadro legislativo o regolamentare, o dello sviluppo, della verifica o dell'utilizzo di tali strumenti e servizi.

Tenuto conto delle risultanze dello Studio preliminare allegato alla Carta e delle modalità di utilizzo dell'intelligenza artificiale diffuse in ambito giudiziario, tra le quali è compresa anche l'analisi predittiva, si stabilisce che, il primo principio da applicare è quello del rispetto dei diritti fondamentali.

In secondo luogo, in base al principio di non discriminazione, gli strumenti di intelligenza artificiale impiegati in ambito giudiziario devono prevenire specificamente lo sviluppo o l'intensificazione di discriminazioni tra persone o gruppi di persone.

Nel trattamento di decisioni e dati giudiziari, poi, il principio di qualità e sicurezza impone di utilizzare fonti certificate e dati intangibili con modelli elaborati multi disciplinarmente, in un ambiente tecnologico sicuro. In base al principio di trasparenza, imparzialità ed equità è necessario rendere accessibili e comprensibili le metodologie di trattamento dei dati, autorizzando verifiche esterne. Infine, occorre assicurare il controllo dei sistemi di intelligenza artificiale agli utilizzatori, che devono essere attori informati, consapevoli delle proprie scelte.

L'idea di fondo che emerge dalla Carta etica, dunque, è che l'utilizzo dell'intelligenza artificiale, deve essere incoraggiata per favorire la prevedibilità e l'uniformità degli orientamenti giurisprudenziali e, quindi, la certezza del diritto. Tuttavia, tale tecnologia deve essere solo uno strumento d'ausilio del giudice, che non può e non deve essere sostituito. L'osservanza dei citati principi deve essere assicurata sin dalla fase di

[153] V. C. BARBARO, *Cepej, adottata la prima Carta etica europea sull'uso dell'intelligenza artificiale (AI) nei sistemi giudiziari*, in *Questione giustizia on line*, 7 dicembre 2018, www.questionegiustizia.it/articolo/cepej-adottata-la-prima-carta-etica-europea-sull-uso-dellintelligenza-artificiale-ai-nei-sistemi-giudiziari_07-12-2018.ph

progettazione e di apprendimento del sistema, secondo un approccio "*ethical-by-design*" o "*human-rights-by-design*", e richiede la vigilanza di un'autorità indipendente, munita di poteri di certificazione e di controllo. La Carta contiene in appendice una raccomandazione sull'uso delle tecnologie di intelligenza artificiale, un glossario e uno strumento di autovalutazione, che permette una verifica sul livello di adesione ai principi fondamentali in essa enunciati.

Anche l'UNESCO, sulla base dei risultati del sondaggio svolto nel 2023 in precedenza citato, ha elaborato una bozza di linee guida specifiche[154], per l'utilizzo di sistemi di intelligenza artificiale, inclusa quella generativa, da parte delle istituzioni giudiziarie, dei magistrati e delle Corti, aperta alla consultazione pubblica fino al 5 settembre 2024, con la versione definitiva prevista per novembre 2024.

La bozza rientra nel programma "*AI and the Rule of Law*" della Global Judges Initiative e costituisce una guida completa per garantire che l'impiego dell'intelligenza artificiale sia conforme ai principi fondamentali della giustizia, dei diritti umani e dello Stato di diritto e per prevenire potenziali abusi o errori derivanti da decisioni automatizzate, proteggendo l'integrità del processo giudiziario e fornendo un quadro per la formazione dei giudici. Le linee guida prevedono principi generali, validi, sia per le organizzazioni e le istituzioni responsabili dell'amministrazione della giustizia, sia per gli operatori, principi di governance, destinati alle istituzioni responsabili e principi di utilizzo, destinati al singolo magistrato.

A differenza di quelle finora esaminate, le norme dell'*AI Act*, come si è detto, sono giuridicamente vincolanti e, quindi, rappresentano una svolta epocale nella regolamentazione delle nuove tecnologie connesse all'intelligenza artificiale, anche con riferimento specifico al settore giudiziario.

A tal proposito occorre rilevare che già l'art.22 del Regolamento 2016/679 sulla tutela della privacy (GDPR) che disciplina il «*Processo decisionale automatizzato relativo alle persone fisiche, compresa la profilazione*» stabilisce che «*L'interessato ha il diritto di non essere sottoposto a una decisione basata unicamente sul trattamento automatizzato, compresa la profilazione*[155]*, che produca effetti giuridici che lo*

[154] UNESCO, *Global toolkit on AI and the rule of law for the judiciary*, 2023, https://unesdoc.unesco.org/ark:/48223/pf0000387331?posInSet=1&queryId=347c2561-bea8-44c8-89f6-d0be80b17064

[155] L'articolo 4 del GDPR definisce la profilazione come «*qualsiasi forma di trattamento automatizzato di dati personali consistente nell'utilizzo di tali dati personali per valutare determinati aspetti personali*

riguardano o che incida in modo analogo significativamente sulla sua persona.» Il Gruppo di lavoro[156] oggi sostituito dal *European data protection board* ha precisato che il «*processo decisionale esclusivamente automatizzato*» cui fa riferimento l'art.22 del GDPR consiste nella capacità di prendere decisioni impiegando mezzi tecnologici senza alcun coinvolgimento umano. Una decisione automatizzata può essere presa senza aver creato un profilo dell'individuo oppure può richiedere la profilazione, a seconda del modo in cui i dati vengono utilizzati[157].

In entrambi i casi, quando la decisione è assunta esclusivamente da una macchina si verifica una fattispecie estrema che, ai fini della tutela della riservatezza dei dati, è vietata dalla norma dell'art.22 GDPR.

Il paragrafo successivo della norma in esame, tuttavia, introduce una serie di eccezioni, stabilendo che il divieto di decisioni unicamente automatizzate «*non si applica nel caso in cui la decisione: a) sia necessaria per la conclusione o l'esecuzione di un contratto tra l'interessato e un titolare del trattamento; b) sia autorizzata dal diritto dell'Unione o dello Stato membro cui è soggetto il titolare del trattamento, che precisa altresì misure adeguate a tutela dei diritti, delle libertà e dei legittimi interessi dell'interessato; c) si basi sul consenso esplicito dell'interessato*.» Pertanto, in base a tale norma l'eventuale automazione integrale di una procedura giudiziaria potrebbe essere legittimata dal consenso dell'interessato e dall'adozione, da parte dei singoli Stati membri, di specifiche leggi che prevedano misure adeguate a tutela dei diritti e delle libertà degli interessati.

Come rilevato in dottrina[158], dunque, l'effettività delle tutele previste dall'art. 22 GDPR, basate sull' esclusività del fattore automazione, tende ad essere confinata in un perimetro sfuggente. Peraltro, la definizione di decisione integralmente automatizzata è

---

relativi a una persona fisica, in particolare per analizzare o prevedere aspetti riguardanti il rendimento professionale, la situazione economica, la salute, le preferenze personali, gli interessi, l'affidabilità, il comportamento, l'ubicazione o gli spostamenti di detta persona fisica».

[156] Gruppo di lavoro articolo 29 dei Garanti Privacy europei *Linee Guida su profilazione e processi decisionali automatizzati*, www.privacy.it

[157] Il Gruppo di lavoro, per esplicitare il concetto, fa il seguente esempio: una multa per eccesso di velocità rilevata sulla base delle prove provenienti da telecamere è una decisione automatizzata che non coinvolge profili. Si tratterebbe invece di profilazione se l'importo della multa fosse il risultato di una valutazione che coinvolge altri fattori quali le abitudini di guida, altre violazioni al codice della strada, ecc.

[158] G. FASANO, *L'interpretazione estensiva della nozione di "decisione automatizzata" ad opera della Corte di giustizia: una prospettiva più ampia ma ancora fragili tutele per le libertà fondamentali*, in *Rivista italiana di informatica e diritto*, 2024, p.6.

incerta. Inoltre, occorre considerare che gli algoritmi, anche quando la decisione non è basata unicamente sul trattamento automatizzato di aspetti personali, possono prevalere, di fatto, sulla valutazione dell'uomo.

Ferma restando la disciplina dettata dal GDPR, l'*AI Act*, in relazione al settore giudiziario, fa riferimento specifico ai sistemi di intelligenza artificiale destinati a essere usati da un'autorità giudiziaria o per suo conto, per assisterla nella ricerca e nell'interpretazione dei fatti e del diritto e nell'applicazione della legge a una serie concreta di fatti, o per la risoluzione alternativa delle controversie. Tali sistemi sono classificati "ad alto rischio" e sottoposti alla relativa disciplina, in considerazione del loro impatto potenzialmente significativo sulla democrazia, sullo Stato di diritto, sulle libertà individuali e sul diritto a un ricorso effettivo a un giudice imparziale.

Nel considerando 61 dell'*AI Act* si precisa che non sono ad alto rischio i sistemi destinati ad attività amministrative puramente accessorie, che non incidono sull'effettiva amministrazione della giustizia nei singoli casi, quali l'anonimizzazione o la pseudonimizzazione di decisioni, documenti o dati giudiziari, la comunicazione tra il personale, i compiti amministrativi. Si tratta di un ventaglio di attività molto ampio che riguarda una molteplicità di compiti, nella maggior parte dei Paesi, affidati a personale diverso dai magistrati. In Italia, invece, tali attività che, peraltro, comportano un grosso dispendio di tempo e notevoli problemi organizzativi, sono svolte dai magistrati. Pertanto, l'impiego dell'intelligenza artificiale potrebbe rivelarsi particolarmente utile.

Il considerando 61 puntualizza, altresì, che l'utilizzo di strumenti di intelligenza artificiale può fornire sostegno al potere decisionale dei giudici o all'indipendenza del potere giudiziario, ma non deve sostituirlo: «*il processo decisionale finale deve rimanere un'attività a guida umana*.»

Secondo parte della dottrina[159], le disposizioni finora esaminate suscitano qualche dubbio. Infatti, in primo luogo la previsione dell'utilizzo di applicazioni di IA ritenute ad alto rischio per «*l'assistenza di un'autorità giudiziaria*», è eccessivamente generica e non chiarisce cosa debba intendersi per «*autorità giudiziaria*».

Se si optasse per un'interpretazione restrittiva di tale espressione si potrebbe circoscrivere la nozione, all'«*autorità in grado di fornire l'effettiva tutela*

[159]A. CORRERA, *Il ruolo dell'Intelligenza artificiale nel paradigma europeo dell'E-justice. Prime riflessioni alla luce dell'AI Act*, in *Quaderni AISDUE*, 2, 2024

giurisdizionale» contenuta nell'art. 47 della Carta dei diritti fondamentali. Tuttavia, in tal modo, il lavoro di una Procura della Repubblica, ad esempio, non rientrerebbe nell'ambito di applicazione delle norme.

Inoltre, si rileva che lo stesso sistema di reperimento di informazioni legali o di ricerca della giurisprudenza, quando è utilizzato nella pratica privata (ad esempio, dagli avvocati), non è considerato, almeno in teoria, ad alto rischio. Pertanto, secondo l' *AI Act* la diversa natura (privata) dell'operatore del medesimo settore giudiziario, può cambiare la categoria di rischio del sistema di IA utilizzato. Inoltre, se un sistema di IA, commercializzato esclusivamente per la pratica privata, seppure involontariamente fosse utilizzato da un giudice, non sarebbe considerato ad alto rischio.

Infine, la dottrina in esame rileva che, nonostante il Regolamento prescriva, per l'addestramento delle macchine intelligenti, la necessità di utilizzare set di dati accessibili, sicuri, privi di errori e pregiudizi, non indica quali azioni potranno essere realizzate in concreto per superare le criticità connesse al fenomeno delle black box e garantire che sia rispettato il dovere del giudice di motivare le sue decisioni, il principio di indipendenza e di imparzialità del giudice e il diritto ad una tutela giurisdizionale effettiva.

Pertanto, si ritiene sia necessario individuare i possibili sviluppi del sistema giudiziario e identificare modelli compatibili con il rispetto dei diritti fondamentali.

In assenza di strumenti che aiutino il giudice e il personale giudiziario ad acquisire un'adeguata formazione digitale e ad accrescere la consapevolezza dei rischi connessi all'impiego delle nuove tecnologie, è difficile che essi siano in grado di discostarsi dalla decisione o dai suggerimenti dell'IA.

Il d.d.l. AS 1146 in discussione al Senato stabilisce che l'intelligenza artificiale potrà essere usata solo per attività di supporto quali l'organizzazione e la semplificazione del lavoro oppure per la ricerca giurisprudenziale e dottrinale. Il legislatore precisa, altresì, che è riservata al magistrato la decisione sull'interpretazione della legge, la valutazione dei fatti e delle prove e sull'adozione di ogni provvedimento, inclusa la sentenza.

La disciplina giuridica, dunque, con riferimento all'attività di *iuris dictio* intende tenere distinte la mente meccanica e quella umana, assegnando alla prima una posizione accessoria e di mero ausilio rispetto alla seconda. Inoltre, vuole salvaguardare la libertà del giudice di decidere se ricorrere o meno al supporto automatico.

Alla base dell'*AI Act* c'è la convinzione che la giurisdizione non sia riducibile a un mero calcolo matematico e richieda una valutazione cognitiva che solo la mente umana è al momento in grado di compiere, confermando la centralità dell'uomo rispetto all' ecosistema digitale[160].

Tale approccio impone di usare l'intelligenza artificiale anche in ambito giudiziario, in modo appropriato senza essere usati da essa. Per rapportarsi con una società dove l'intelligenza artificiale eserciterà una influenza paragonabile, ma non identica, a quella dell'intelligenza umana, il mondo della giustizia deve modernizzare tutta la sua visione della realtà. La giustizia deve avvalersi in un modo innovativo, del linguaggio dei numeri e degli algoritmi, inserendoli, tuttavia, in un contesto di valori e di umanità.

Peraltro, la Costituzione lega la giurisdizione alla persona umana e conferma che l'intelligenza artificiale nel settore giudiziario può assolvere solo ad una funzione servente e consulenziale.

L'art. 102 Cost., infatti, fa chiaramente riferimento al giudice persona fisica (quale giudice naturale precostituito per legge) e richiede che la funzione giurisdizionale sia affidata a magistrati, l'art. 111 Cost. per garantire il giusto processo fa riferimento ad un giudice terzo e imparziale, così come l'art. 101 Cost. vuole che i giudici siano soggetti solo alla legge, in tal modo escludendo che possano essere vincolati all'esito dell'algoritmo predittivo. L'art. 24 Cost., infine, postula il pieno dispiegamento del diritto di difesa tra esseri umani che sarebbe di fatto compresso di fronte alla potenza degli algoritmi capaci di processare decenni e decenni di pronunce giurisprudenziali.

Peraltro, anche la giurisprudenza costituzionale ha avuto occasione di dichiarare l'incostituzionalità dei c.d. automatismi decisionali, in varie occasioni[161]. In particolare,

---

[160] G. DE MINICO, *Giustizia e intelligenza artificiale: un equilibrio mutevole*, in *Rivista Associazione Italiana dei Costituzionalisti*, 2024, 2 p.108; A. BALSAMO, *L'impatto dell'intelligenza artificiale nel settore della giustizia*, cit., p.3; S. ARDUINI, *La "scatola nera" della decisione giudiziaria: tra giudizio umano e giudizio algoritmico*, in *BioLaw Journal – Rivista di BioDiritto*, 2021, 2, pp.453.

[161] Si consideri la giurisprudenza costituzionale che, in materia di adozione, si è opposta alla previsione di una differenza di età fissa tra adottato e adottanti (Cfr. Corte cost. n. 303 del 1966; n. 140 del 1990; n. 148 del 1992; n. 283 del 1999) e le sentenze in cui la Corte costituzionale ha riscritto la norma nel senso di consentirne un'applicazione rispondente al superiore interesse del minore (cfr. Corte cost. n. 31 del 2012; n. 7 del 2013). Si considerino, altresì, le decisioni con cui la Corte costituzionale ha censurato la presunzione assoluta di adeguatezza della misura della custodia cautelare in carcere in relazione ad alcune categorie di delitti (Cfr. Corte cost. n. 265 del 2010; n. 128, n.164, n. 231, n. 331 del 2011; n. 110 del 2012; n. 57, n. 213, n. 232 del 2013).inoltre, si può richiamare, la decisione costituzionale n. 151 del 2009 che, in tema di fecondazione medicalmente assistita di tipo omologo, ha censurato l'applicazione rigida della norma della legge n. 40 del 2004, Norme in materia di procreazione medicalmente assistita, che

nelle presunzioni algoritmiche, come nelle presunzioni assolute vi è un legame statistico tra il fatto ignoto e il fatto noto. I meccanismi presuntivi, in linea generale, sono potenzialmente incostituzionali in quanto possono entrare in conflitto con il principio di uguaglianza rischiando di parificare situazioni oggettivamente diverse.

Specialmente nel diritto penale, poi, l'uso degli automatismi non consente di valutare determinate circostanze individuali che sono altamente condizionate dalle caratteristiche e peculiarità di ciascuna fattispecie.

L'acclarata incostituzionalità delle presunzioni e, in generale di tutti gli automatismi decisionali, soprattutto nel campo del diritto penale, restituisce ai giudici la facoltà di decidere sul singolo caso conferendo loro la delega del bilanciamento. Con tale strumento, come rilevato da autorevole dottrina[162] la Corte elimina una graduatoria rigida degli interessi in gioco e affida al giudice il compito di individuare, caso per caso, il punto del loro equilibrio.

Ispirandosi a tali principi l'Unione delle Camere penali italiane[163], il 17 gennaio 2025, ha presentato il documento "*Giustizia penale, scienza e intelligenza artificiale. La carta dei valori dell'Unione delle camere penali italiane*", una dichiarazione di impegno "politico" nel solco del giusto processo redatto dall'Osservatorio Scienza Processo e Intelligenza artificiale.

L'Osservatorio premette che sono attualmente disponibili sistemi di AI di diversa natura (sistemi esperti, machine learning) che possono operare in tutte le fasi del processo penale ed essere usati sia dai giudici, che dai consulenti tecnici e i periti. Tali strumenti, se non adeguatamente controllati potrebbero creare squilibrio nell'accertamento della verità processuale.

Inoltre, tali tecnologie rendono più complesso il rapporto uomo-macchina, tenendo conto dei reciproci bias, cognitivi per il primo e potenzialmente discriminatori per la seconda.

---

limitava a tre il numero massimo di embrioni fecondabili e destinati ad un unico e contemporaneo impianto nell'utero della donna.

[162] R. BIN, *Diritti e argomenti. Il bilanciamento degli interessi nella giurisprudenza costituzionale*, Giuffrè, Milano, 1992, pp. 91-92.

[163] *Giustizia penale, scienza e intelligenza artificiale. La carta dei valori dell'Unione delle camere penali italiane*, camerepenali.it

L’Unione intende promuovere un approccio multidisciplinare alle tecnologie, consapevole del grande valore del metodo scientifico applicato al processo.

Tuttavia, tale metodo deve essere servente rispetto, sia ai principi generali in materia di IA (beneficenza, non maleficenza, autonomia, trasparenza, responsabilità e giustizia), sia ai principi del giusto processo, sanciti dalla Costituzione.

Pertanto, è necessario garantire la presunzione di innocenza, il diritto al contraddittorio e la possibilità di confutare le prove, la motivazione delle decisioni e il controllo della loro legalità e logicità nonché il principio “*in dubio pro reo”*.

Inoltre, occorre promuovere la conoscenza degli errori e delle distorsioni cognitive, umani e tecnologici, di tutte le parti del processo, compresi i periti e, quindi, promuovere l’applicazione di modelli e procedure per minimizzare il rischio di errori giudiziari e di modelli e procedure che assicurino la scientificità dei metodi adottati da consulenti e periti.

Infine, l’Unione si impegna a promuovere e verificare l’indipendenza di chi sviluppa algoritmi, la messa a punto di procedure controllate per l’acquisizione dei dati, evitando discriminazioni; a sollecitare la cybersicurezza per garantire affidabilità e sicurezza dei sistemi informatici e la trasparenza degli algoritmi utilizzati in ambito penale e, come corollari di tale principio, la spiegabilità dei modelli, la loro possibile ispezione e la dimostrabilità degli esiti.

Si propone, altresì, l’introduzione del divieto di decisioni completamente automatizzate nel giudizio penale e di procedure chiare e precise nella delega di funzioni ai sistemi informatici rispetto a tutte le fasi del processo decisionale, a partire dall’acquisizione delle informazioni, fino alla decisione e relativa esecuzione.

CAPITOLO III

La decisione robotica

Sommario: 3.1 La decisione robotica. - 3.1.1. La relazione tra esseri umani e macchine: umanesimo, postumanesimo e transumanesimo. -3.2. Il tecnoumanesimo. -3.3 La funzione delle norme.

3.1 La relazione tra esseri umani e macchine.

L'applicazione dell'intelligenza artificiale in ambito giudiziario ha consentito di creare dei sistemi predittivi che permettono di analizzare grandi quantità di dati per trattare un caso, contestualizzarlo secondo specifiche caratteristiche e anticipare le decisioni che potrebbero essere prese. Attraverso sistemi di giustizia predittiva, dunque, l'intelligenza artificiale aiuta gli avvocati e le parti a valutare le possibilità di successo di una controversia, prevedendone il risultato finale e supporta il magistrato giudicante, suggerendogli la decisione più coerente.

I rapidi progressi della tecnica potrebbero condurre, in un futuro non molto lontano, alla realizzazione di sistemi di intelligenza artificiale che non si limitano a fornire solo un supporto agli operatori del diritto. Si intende far riferimento a sistemi di intelligenza artificiale cosiddetta forte (AGI) che potrebbero essere in grado di raggiungere stati cognitivi assimilabili agli stati mentali umani o persino ad una super intelligenza artificiale (ASI) capace di superare le abilità cognitive umane.

L'applicazione di tali sistemi in ambito giudiziario consentirebbe di sostituire integralmente il giudice con il robot, la decisione umana con quella robotica. Quando si discute di decisione robotica integralmente automatizzata, dunque, non si fa riferimento a strumenti informatici che sono un mero supporto per la comunicazione, la gestione dei documenti legali o la ricerca di informazioni, ma a macchine, i cosiddetti robogiudici, che assumono un ruolo esclusivo nel prendere decisioni in contesti legali complessi. Pertanto, si fa riferimento ad un processo decisionale automatizzato nel quale l'essere umano non esercita alcun controllo funzionale, delegando la decisione all'algoritmo[164].

[164] S. SAPIENZA, *Decisioni algoritmiche e diritto,* Giuffrè Milano, 2024, p.36.L'autore individua quattro diverse classi di processi decisionali che comprendono il processo decisionale analgico, in cui non è coinvolto un sistema informatico e il processo decisionale aritmetico nel quale l'algoritmo fornisce supporto a fini computazionali eseguendo solo le istruzioni dell'uomo che può interrompere la macchina e sostituirsi ad essa (ad esempio la calcolatrice). Vi sono, poi, processi decisionali automatici nei quali l'essere umano fornisce istruzioni all'algoritmo che, tuttavia, da un contributo imprescindibile ai fini della

I giudici robot potrebbero essere pensati come sistemi cognitivi artificiali altamente specializzati, in grado di gestire in modo automatizzato una vasta gamma di casi legali in parallelo, anziché in serie come avviene spesso con i giudici umani. La capacità di lavorare contemporaneamente su tutti i casi pendenti potrebbe ridurre i tempi e i costi del processo rendendo il sistema più accessibile ed efficiente per tutti. Allo stesso tempo, la corretta configurazione dei sistemi in esame consentirebbe di garantire una giustizia che, oltre ad essere veloce, potrebbe essere giusta e corrispondente alla verità. Inoltre, la decisione robotica potrebbe contribuire ad assicurare una maggiore imparzialità, uniformità, oggettività e certezza nelle decisioni giuridiche, non più orientate ai valori e, quindi, discrezionali. Del resto, l'esigenza di certezza non è nuova e già durante l'Illuminismo aveva portato alla nascita dei codici e alla proceduralizzazione del diritto e del processo[165]. Peraltro, nella prospettiva Illuministica il giudice era considerato come una macchina, "*bouche de la loi*". Pertanto, la sua decisione doveva essere impersonale e si basava sull'applicazione del sillogismo[166].

Tuttavia, tale visione si scontra con una realtà processuale che non si presta al riduzionismo sillogistico, superato dalla constatazione del ruolo che l'interpretazione e la discrezionalità giudiziale ricoprono nell'opera di applicazione del diritto ai casi concreti. All'estremo opposto rispetto alla matematica processuale del giudice "*bocca della legge*", dunque, si colloca la prospettiva elaborata da alcuni esponenti del realismo giuridico americano[167] secondo cui «*la giustizia è ciò che il giudice ha mangiato a*

---

decisione (ad esempio l'autovelox). Il processo decisionale assistito, invece, è caratterizzato dal fatto che l'essere umano porta a termine la decisione finale alla quale l'algoritmo ha contribuito (ad esempio i sistemi utilizzati per filtrare i curriculum vitae, nella selezione del personale). Infine, vi è il processo decisionale automatizzato nel quale l'algoritmo sostituisce integralmente il decisore umano (ad esempio i sistemi di credit scoring con i quali le banche analizzano l'indice di affidabilità dei potenziali clienti). Tali tipologie di processi decisionali, fatto salvo quello analogico, si avvalgono dell'algoritmo che interviene nella decisione con un grado di penetrazione differente che va dal mero supporto alla delegazione.

[165] F. ROMEO, *Algoritmi di giustizia ed equità nel diritto. Quando razionalità ed emozionalità convergono*, in *I-lex. Scienze Giuridiche, Scienze Cognitive e Intelligenza Artificiale*, www.i-lex.it, 2021, I.

[166] S. ARDUINI, *La "scatola nera" della decisione giudiziaria: tra giudizio umano e giudizio algoritmico*, in *BioLaw Journal – Rivista di BioDiritto*, 2021, 2, pp.453-470.

[167] Il realismo giuridico americano si sviluppò negli Stati Uniti d'America negli anni '20 e '30 del Novecento, con l'opera di alcuni giuristi i quali ispirandosi al positivismo in senso filosofico, al pragmatismo americano e all'empirismo logico sostengono che esiste un solo universo governato da leggi naturali (nel senso in cui sono leggi naturali le leggi della fisica – per es. gravitazione universale di Newton) e che il metodo scientifico è l'unico che permette di conoscere davvero la realtà. Questo metodo, quindi, dovrebbe essere usato da tutti, anche da filosofi e giuristi.

colazione»[168]. In tali prospettive, la decisione giudiziale non è una mera operazione logica ed è influenzata da una componente intuitiva e dalla personalità del giudice, che non è più ridotto ad una sorta di automa applicatore della legge. Pertanto, nella decisione assumono rilevanza i fattori sociali e culturali e quelli psicologico-emotivi nonché i valori e i significati di cui il giudice è veicolo.

Il processo, dunque, è costantemente animato dalla tensione fra la proceduralizzazione disumanizzante e i rischi dell'arbitrio della soggettività.

3.1.1 La relazione tra esseri umani e macchine: umanesimo, postumanesimo e transumanesimo.

L'ipotesi dell'introduzione di un decisore algoritmico che si sostituisca al decisore umano suscita reazioni contrastanti che riflettono l'evoluzione del pensiero filosofico relativo al rapporto tra l'uomo e le macchine.

L'essere umano è stato storicamente definito cercando di stabilire cosa lo differenzi in modo costitutivo da tutto ciò che è diverso da sé. Secondo la prospettiva dell'Umanesimo[169] l'identità umana è definita e definitiva. La specie umana nella sua varietà antropologica è accomunata da bisogni, aspirazioni e ideali che derivano soprattutto dalla razionalità che la caratterizza. Secondo tale impostazione, l'essere

[168] J.FRANK.*Courts on Trial, Myth and Reality in American Justice*, 1949, p.405; ID., *Are Judges Human? Part 1: The Effect on Legal Thinking of the Assumption That Judges Behave Like Human Beings*, in *University of Pennsylvania Law Review*, 80, 1, 1931, pp.17 ss.

[169] Il termine "umanesimo" è impiegato per indicare il movimento culturale sviluppatosi in Italia nel corso del XV e XVI secolo basato sulla riscoperta dei classici latini e greci nella loro storicità. L'obiettivo di tale movimento è recuperare dai classici, conoscenze e usanze da utilizzare nella quotidianità per avviare una "rinascita" della cultura europea dopo il Medioevo. Il medesimo termine è impiegato anche in un'accezione che ha le sue radici in Terenzio, il commediografo romano, il quale, in una formula ripresa da Cicerone, esprime l'essenza dell'Umanesimo: «*Homo sum, humani nihil a me alienum puto*», (sono uomo e non considero niente di umano estraneo a me). Sulla base di tale affermazione sono state elaborate le posizioni filosofiche umaniste, che, cioè, mettono l'uomo al centro della realtà. Il termine fu ripreso nell'800 da Ludwig Feuerbach, esponente della Sinistra hegeliana, per esporre le proprie considerazioni filosofiche e, nel corso del Novecento da alcuni intellettuali, legati all'esistenzialismo come Jean-Paul Sartre, Martin Heidegger; Jacques Maritain, Ernst Bloch, Rodolfo Mondolfo e Herbert Marcuse. In particolare, Martin Heidegger nel saggio *La questione della tecnica*, affermò che la tecnica consiste nella produzione di un artefatto attraverso il quale l'uomo disvela la verità. Tuttavia, la tecnica moderna, con il suo carattere marcatamente industriale, concepisce la natura e, dunque, anche l'uomo, in termini di sfruttamento, trasformandolo in una riserva di materiali da impiegare, che vengono ridotti a "fondo", "risorsa". Pertanto, l'uomo, da artefice dello sfruttamento rischia di trasformarsi in vittima di questa stessa sua azione. Secondo la tesi in esame, infatti, l'uomo contemporaneo, non è in grado di esercitare un controllo sulla tecnica ed è indotto a trattare il reale come una risorsa consumabile. Il rimedio, dunque, secondo Heidegger, sta nell'abbandono delle cose e del potere della tecnica per reintrodurre la possibilità di scelta. Ciò restituisce all'uomo una condizione di trascendenza rispetto agli enti, e rende la tecnica non più una necessità indiscutibile, ma un problema.

umano razionale non ha alcun rapporto e non può in alcun modo essere confrontato con la tecnica, con il mondo animale e con l'ambiente. Esso, infatti, rappresenta il vertice di tutta la catena naturale che collega gli esseri viventi, si eleva a giudice e a misura delle cose, e dei coinquilini che si ritrova nel Pianeta su cui abita. Nella prospettiva umanistica, dunque, l'uomo è al centro del mondo, è autonomo rispetto ad esso e si pensa separato, non solo dall'ambiente che lo circonda, ma anche dai mezzi che lo collegano a quest'ultimo, tra i quali vi è la tecnica o tecnologia, che è considerata un mero accessorio, utilizzato a intermittenza e dal quale vi è necessità di allontanamento, terminato l'uso[170]. Nella prospettiva umanistica, infatti, la *techne* è un atto compensativo, vale a dire una sorta di stampella che l'uomo si procura per potenziare i propri predicati senza metterli in discussione. Essa provoca ribrezzo in quanto è considerata invasiva e innaturale, come il più letale degli allergeni che l'uomo possa incontrare sul suo cammino[171].

La prospettiva umanista, tuttavia, è stata progressivamente superata.

I primi germi del cambiamento si manifestarono già con l'avvento della cibernetica, disciplina che, come si è detto[172], il matematico Norbert Wiener negli anni '40 del 1900 elaborò per studiare i meccanismi dell'autoregolazione e del controllo presenti negli organismi viventi e replicarli nelle macchine. Secondo il matematico la creazione di macchine con caratteristiche simili a quelle umane e, quindi, l'ibridazione tra l'uomo e le macchine, avrebbe permesso di liberare l'uomo dalla sua condizione lavorativa robotica, lasciando i lavori meccanici alle macchine stesse, e di sfruttare in modo più intelligente le grandi capacità mentali umane, penalizzate da società non capaci di trarne appieno benefici.

Nel 1977 lo studioso egiziano Ihab Hassan, per definire la condizione degli esseri umani, utilizzò per la prima volta l'espressione "*post-human"* nel testo *Prometheus as Performer*: *Toward a Posthumanist Culture?* in cui egli sosteneva che l'epoca a lui contemporanea, alla luce delle imponenti rivoluzioni tecnologiche e dell'imporsi culturale del postmoderno, esigesse un radicale cambiamento della forma umana,

[170] I. DE DOMINICIS, *L'interiorizzazione della tecnica. Dal protesico all'estetico*, in *Lo Sguardo - rivista di filosofia*
2017, 24, p. 128
[171] R. MARCHESINI, *Il tramonto dell'uomo. La prospettiva post-umanista*, Dedalo, Bari 2009, p. 20.
[172] Cfr. Supra Cap I p.4

inclusi i desideri e le sue rappresentazioni esterne[173]. Secondo l'autore «*Dobbiamo comprendere che la forma umana — incluso il desiderio umano e ogni rappresentazione esterna — deve cambiare radicalmente e pertanto deve essere riconcepita [...l cinquecento anni di umanesimo devono incamminarsi verso una fine, di modo che l'umanesimo dia luogo a qualcosa che dobbiamo senza riserve chiamare postumanesimo*.»[174]

Negli Ottanta del 1900, poi, fenomeni come la diffusione della fantascienza[175], l'affermazione dell'informatica e delle biotecnologie che mettono in discussione la specificità dell'umano nonché la diffusione del pensiero darwiniano, portarono al superamento dell'idea tradizionale di essere umano e alla speranza che l'umanità potesse un giorno essere qualcosa di più rispetto a ciò che era fino ad allora.

Negli anni Novanta la corrente postumana si affermò e influenzò in modo trasversale la cultura, emergendo da una comunicazione globale che prese le mosse dalla rivoluzione digitale e dalla messa a punto dei grandi network interattivi[176].

Nel 1992, Jeffrey Deitch utilizzò l'espressione "*post umano*" nel corso di un'omonima mostra dedicata a diversi artisti le cui opere mettevano in discussione la tradizionale identificazione umanistica dell'essere umano[177]. La mostra raccolse le tendenze

[173] F. FERRANDO, *Il Postumanesimo filosofico e le sue alterità*, Edizioni ETS, Pisa, 2016, pp.1 e ss.

[174] IHAB HASSAN, *Prometheus as Performer: Toward a Posthumanist Culture?* in *The Georgia Review* 1977, 31, 4, p.212

[175]La fantascienza è un genere letterario e cinematografico che si fonda su ipotesi o intuizioni di carattere scientifico e li mescola alla fantasia. Il genere ha il suo precursore nel romanzo scientifico dell'Ottocento. che ebbe i suoi massimi esponenti nel francese Jules Verne e nel britannico H.G. Wells, e trova un antesignano, nel Frankenstein di Mary Shelley (1818). Il genere raggiunse la sua età dell'oro a cavallo tra gli anni '40 e '50 del Novecento. Autori come Isaac Asimov, Robert A. Heinlein, Ray Bradbury, sono considerati i padri della fantascienza moderna. La Seconda guerra mondiale e le tensioni atomiche portarono il genere da un generale clima positivo e ottimistico, che aleggiava nella fantascienza della prima metà del Novecento, ad un atteggiamento disfattista e pessimistico. Nel dopoguerra, infatti, si diffuse la corrente distopica cui appartenevano romanzi e pellicole che descrivevano una società negativa, anti-utopica, in cui la tecnologia causa il disfacimento sociale. Dalle tensioni della Guerra fredda nacquero, altresì, due sottogeneri: apocalittico e postapocalittico. Negli anni '80, poi, la fantascienza fu influenzata dalla teoria del cyborg o organismo cibernetico o bionico, un essere al confine tra uomo e macchina che rappresenta una nuova soggettività politica, un nuovo modo di pensare la corporeità, trascendere l'umanesimo e dar forma a un uomo in cui vi sia sempre più una commistione tra biologico e tecnologico.

[176] I. DE DOMINICIS, *L'interiorizzazione della tecnica. Dal protesico all'estetico*, cit., p. 128

[177] La mostra ha avuto luogo nel Museo di Arte Contemporanea di Losanna nel 1992, giungendo in Italia nell'ottobre dello stesso anno al Castello di Rivoli, in provincia di Torino. Sono stati esposti i lavori di Kiki Smith, Pia Stadtbäumer, Paul McCarthy, Robert Gober e le opere a esplicito riferimento pornografico di Felix GonzalezTorres, Ashley Bickerton, Charles Ray e Jeff Koon.

artistiche più significative degli anni Ottanta dedicate alla ridefinizione del corpo umano a partire dalle sfide che la tecnologia e la realtà virtuale hanno posto in essere.

Nel descrivere il progetto della mostra, Deitch rilevò che la chirurgia plastica, la ricostruzione genetica e gli innesti di componenti elettronici, diventando prassi comune e, alterando radicalmente la struttura delle interazioni umane, rappresentano le basi di un nuovo stadio evolutivo dell'essere umano che tende a modificare o a perdere le caratteristiche umane, diventando postumano[178].

Nel 1995 con il saggio *The Posthuman Condition*, Robert Pepperell fornì una prima definizione teoretica del postumano, inteso come un progetto di ricerca sull'essere umano basato su nuovi presupposti e in grado di superare la tradizione umanistica[179]. L'Autore sosteneva che non vi è una netta separazione tra l'uomo e la realtà che lo circonda. In particolare, l'intero progresso tecnologico della società umana è diretto verso la trasformazione della razza umana; pertanto, le macchine e, nello specifico, i computer saranno sempre più simili agli umani che, a loro volta si sviluppano per piacere di più ai computer[180].

Mentre l'era umanista è fondata sulla certezza circa il funzionamento dell'universo e la posizione dell'essere umano, dunque, l'era postumana è caratterizzata dall'incertezza. Infatti, nel pensiero postumano l'identità dell'uomo non è definita, in quanto egli stesso non è un essere finito ed è in continuo divenire. In questo divenire l'uomo si ibrida, in quanto essere vivente che muta, inserito in un ambiente. L'ibridazione umana avviene sia con l'ambiente naturale circostante, sia con l'ambiente tecnico.

L'uomo, quindi, non è più ingabbiato in una identità statica o in un corpo che risulta più un ostacolo che uno slancio. Egli è dopo, è *post*, e, nel suo aver superato questa staticità, si ibrida con lo strumento tecnico e fronteggia l'ambiente circostante con i presupposti per una nuova interpretazione della sua relazione con esso[181].

Ciò che deve essere messo in discussione per i postumanisti non è l'uomo in quanto tale, ma l'immagine dell'uomo prodotta dall'antropocentrismo umanista. Quest'ultimo, difatti, storicamente si è fondato su un sentimento di supremazia ontologica, etica

[178] R. MARCHESINI, *Possiamo parlare di una filosofia postumanista?* in *Lo Sguardo - rivista di filosofia*, 2017, 24, p. 28
[179] R. PEPPERED, *The Posthuman Manifesto*, in *Kritikos*, 2005, n. 2, pp. 1-16.
[180] R. MARCHESINI, *Possiamo parlare di una filosofia postumanista?* cit., p. 29
[181] R. MARCHESINI, *Ivi*, p. 30

dell'umano sul non umano ed ha incentivato la diffusione dell'idea che sia possibile effettuare una ricognizione sull'umano solamente a partire dall'uomo stesso.

Con il postumanesimo, invece, l'umano perde la totale preminenza ontologica, etica sul non umano e viene interpretato come un prodotto storico mutevole, determinato da molteplici pratiche di coniugazione con il non umano.

Nella prospettiva post umanista, dunque, la tecnologia e le macchine che rappresentano il non umano non sono meri accessori, strumenti addizionali di cui avere paura, ma sono una parte integrante della struttura fisica ed intellettiva dell'uomo, che non va temuta[182].

La decostruzione della concezione umanistica tradizionale dell'essere umano costituisce il presupposto essenziale anche per il transumanesimo, una forma di speculazione futurologica elaborata da quelle correnti di pensiero che, considerati i limiti dell'essere umano, sostengono la necessità di potenziarlo e perfezionarlo nei suoi aspetti estetici, anatomici, cognitivi, emotivi, genetici[183].

Secondo quanto sostenuto da studiosi di informatica e di intelligenza artificiale si prefigura una nuova era evoluzionistica post-darwiniana, guidata dalla specie umana stessa, nella quale si avrà una convergenza evolutiva fra uomini e macchine, il funzionamento del cervello sarà ridotto al funzionamento del computer e la mente sarà digitalizzata, in modo da raggiungere una forma di immortalità corporea[184].

In futuro, dunque, si realizzerà un'alterazione radicale della natura dell'uomo, mettendo in correlazione il corpo (materia organica) con i computer (materia inorganica), sino alla totale artificializzazione dell'umano, sostituendo corpo e mente con sussidi meccanici ed informatici.

Il transumanesimo, quindi, a partire da una concezione edonistica dell'essere, dalla negazione ontologica della natura e dal misconoscimento della peculiarità del biologico,

---

[182] Il filosofo Pierre Teilhard de Chardin, nell'opera *Verso la convergenza. L'attivazione dell'energia nell'umanità*, tr. it., Il Segno dei Gabrielli, Verona 2004, p. 75 ad esempio scrive: «*La tecnica ha un ruolo biologico propriamente detto: entra in pieno diritto nel naturale. Da questo punto di vista [...] svanisce l'opposizione tra artificiale e naturale, tra tecnica e vita, perché tutti gli organismi sono il risultato di invenzioni; se una differenza esiste, è in favore dell'artificiale*».

[183] F. FERRANDO, *Postumanesimo, transumanesimo, antiumanesimo, metaumanesimo e nuovo materialismo. Relazioni e differenze*, in Lo Sguardo - rivista di filosofia, 2017, 24, pp.1 e ss.

[184] D.HILL, E. DREXLER, M. MINSKY, H. MORAVEC, *A History of Transhumanist Thought*, in *Journal of Evolution and Technology*, 14, 1, 2005, pp. 1-25; Id., *Human Genetic Enhancements: A Transhumanist Perspective*, in *Journal of Value Enquiry*, 37, 2003, pp. 493-506; R. BRAIDOTTI, *Il postumano. La vita oltre il sé, oltre la specie, oltre la morte*, DeriveApprodi, Roma 2014; T. TOSOLINI, *L'uomo oltre l'uomo. Per una critica teologica a Transumanesimo e Post-umano*, EDB, Bologna 2015.

promuove, eticamente, la transizione verso il trans-umano, abbandonando progressivamente l'umano e la stessa specie umana. L'abbandono del biologico e la transizione verso il virtuale/artificiale/digitale è finalizzato ad espandere le capacità umane, per avere vite migliori e menti migliori, anche attraverso la cancellazione della condizione umana stessa, percepita e vissuta come un limite.

Come si è detto, le diverse tesi filosofiche sul rapporto tra l'uomo e la tecnologia finora esaminate costituiscono la base per affrontare la tematica della sostituzione della decisione umana con quella robotica, in ambito giudiziario.

Le teorie umanistiche sono alla base di un atteggiamento di diffidenza nei confronti della tecnologia che scaturisce dal timore che l'uomo perda la propria centralità nell'universo. Esse alimentano il rifiuto dell'idea che l'intelligenza artificiale possa essere impiegata nel processo fino al punto da sostituire il giudice con un robot. Come rileva autorevole dottrina[185], il rifiuto della tecnologia si basa sulla paura che la macchina possa sostituire l'uomo, rendendo freddo e meccanico il suo mondo, fungibile la sua forza lavoro e superflue, le sue abilità cognitive. In tal modo l'uomo sarebbe destinato a diventare un esemplare in via di estinzione o una specie protetta.

Con riferimento specifico all'ambito giudiziario gli umanisti ritengono che il giudice robot decida in modo asettico, senza consapevolezza circa l'oggetto del contendere e i valori e gli interessi in gioco. Pertanto, la decisione del giudice robot non sarebbe in grado di affermare i principi che garantiscono i valori, la libertà e la dignità degli uomini.

Alla diffidenza si contrappone, invece, l'entusiasmo per l'idea che un robot possa persino sostituirsi ad un giudice nell'attività decisionale, in modo da risolvere il problema della fallibilità e dell'incertezza del giudizio umano. La complessità dell'età contemporanea che è definita età dell'incertezza, infatti, mette a rischio l'attività del giurista che quotidianamente deve confrontarsi con una enorme quantità di leggi speciali all'esterno del Codice civile, di norme sovranazionali e di decisioni basate sui casi concreti e sui precedenti, che hanno dato vita ad un vero e proprio diritto di fonte giurisprudenziale. Pertanto, la possibilità di fare ricorso a robot capaci di decidere in vece dell'uomo, utilizzando strumenti matematici e statistici, presenterebbe vantaggi evidenti in termini di certezza del diritto.

[185] A. PUNZI, *Judge in the machine, e se fossero le macchine a restituirci l'umanità del giudicare?* in A. CARLEO (a cura di), *La decisione robotica*, Il Mulino, Bologna, 2019, p.319

Tuttavia, tale prospettiva solleva numerose preoccupazioni riguardo alla perdita di umanità nel processo decisionale. Il ruolo del giudice, come si è detto, non è solo quello di applicare la legge in modo meccanico ma anche di esercitare il discernimento, la sensibilità e la comprensione delle sfumature delle situazioni. Le dimensioni emotiva-emozionale, esperienziale, interpretativa, intuitiva e intenzionale-motivazionale, insieme alla autoconsapevolezza, alla libertà e alla relazionalità interpersonale, sono elementi e segni che mostrano la unicità umana, che evidenziano la differenza uomo-macchina. Una differenza che esiste e deve rimanere, di principio, per difendere l'unicità dell'umano.

Pertanto, alcuni considerano l'idea del "giudice robot" come il paradigma di una prospettiva distopica, l'immagine emblematica di un mondo disumanizzato e dominato dalla tecnica. Altri, invece, considerano il "giudice robot" «*un buon paradigma idealtipico da opporre alle egemoni dottrine dell'interpretazione, che hanno legittimato l'arbitrio e il soggettivismo», un «modello», una «prospettiva di orientamento […] cui occorre volgersi non per ragioni ideo-logiche, ma logiche e giuridico-positive: perché occorre identificare un confine fra legis-latio e iuris-dictio, se alla distinzione fra le due funzioni si vuole dare ancora un senso, evitando di affogarle nella rigorosa, ma irrealistica, indistinzione kelseniana della nomopoiesi; perché quel confine è presupposto dalle norme costituzionali che disegnano l'assetto istituzionale della Repubblica*»[186].

3.2. Il tecnoumanesimo.

Tra le posizioni estreme in precedenza sommariamente ricostruite si pongono coloro che affrontano il problema del rapporto tra il giudice e la tecnologia e, nello specifico, l'intelligenza artificiale postulando il superamento sia del rifiuto totale, sia dell'eccessivo entusiasmo[187].

Secondo la tesi in esame, è necessario ed inevitabile prendere atto di una realtà contemporanea nella quale l'uomo pensa e decide in constante interazione con le

[186] M. LUCIANI, *La decisione giudiziaria robotica*, in A. CARLEO (a cura di), *La decisione robotica*, Il Mulino, Bologna, 2019, p.63.
[187] A. PUNZI, *Judge in the machine, e se fossero le macchine a restituirci l'umanità del giudicare?* cit, p.328

macchine, e sviluppare un nuovo modo di intendere le sue capacità cognitive e decisionali in termini di tecnoumanesimo o umanesimo digitale[188].

Il processo decisionale deve essere il frutto di una contaminazione tra giudice e macchina che si determinano reciprocamente. Come rileva autorevole dottrina[189], la macchina non mangia, non dorme, non ha crisi sentimentali, né sbalzi d'umore, né preferenze personali o ideologiche ed esamina sempre tutto il fascicolo. La macchina, valuta e risponde a tutte le domande dell'attore, ed a tutte le eccezioni del convenuto, "processa" tutti gli atti ed ha illimitato accesso a tutte le fonti normative ed a tutti i precedenti giudiziali e le pubblicazioni e dati reperibili online. Tuttavia, la macchina ragiona per inferenza e non per deduzione causale, ed opera con comprensione sintattica ma non semantica, sapendo selezionare i significanti e non i significati. Inoltre, non essendo umana, la macchina è priva sia di empatia che di comprensione emotiva, e comunque di immedesimazione.

La macchina è, dunque, fredda e disumana, ma è "intelligente" e razionale e, in prospettiva, potrà essere ancora più sapiens. Questi sono i suoi pregi e difetti e gli elementi che la differenziano dal giudice.

Nella conflittualità del binomio certezza della macchina-umanità del giudice, nessuno dei due volti della giustizia deve prevalere sull'altro, essendo entrambe dimensioni ineludibili del fenomeno giuridico. Più che in contrapposizione, le due componenti vanno, dunque, rilette nell'ambito di una cornice unitaria ed armonizzante in modo da esaltarle per raggiungere l'obiettivo della tutela dei diritti fondamentali.

Pertanto, la macchina raccoglie ed elabora dati, prospetta soluzioni e monitora il percorso decisionale e motivazionale verificando che non sia inficiato da lacune o incongruenze. La decisione, invece, deve essere elaborata dall'uomo che deve motivare l'atto di cui si assume la responsabilità. In tale modello che potrebbe essere istituzionalizzato, la macchina fornisce un ausilio essenziale per velocizzare il processo decisionale del giudice e, nel contempo, vigila sull'intera formazione del giudizio,

---

[188] S. ARDUINI, *La "scatola nera" della decisione giudiziaria: tra giudizio umano e giudizio algoritmico*, cit., p.469

[189] U. RUFFOLO, *Giustizia predittiva e machina sapiens quale "ausiliario" del giudice umano*, in *Lezioni di Diritto dell'Intelligenza Artificiale*, a cura di Ugo Ruffolo, Giappichelli, Torino, 2021, pp.1 e ss.

diventando una sorta di "*grillo parlante digitale*"[190] che aiuta il decisore umano a rendersi conto con maggiore lucidità del modo in cui sta ragionando e decidendo e, eventualmente, a correggere i propri errori.

In tal modo non è scalfito il primato ed il ruolo del giudice umano, che non viene sottoposto al controllo della macchina. Egli sarebbe semplicemente vincolato a consultare anche un algoritmo capace di coadiuvarlo nel reperimento dei testi normativi e giurisprudenziali e di rappresentargli sia una soluzione informata, sia quella che sarebbe la ragionevole aspettativa dei giudicandi[191]. Il giudice umano, al quale è riservata la motivazione della decisione, sarebbe libero di condividere o meno la soluzione algoritmica dando atto delle ragioni della propria scelta. In tal modo, si potrebbe porre un limite alla giurisprudenza creativa, senza inibire l'adeguamento della norma al caso concreto attraverso un'interpretazione evolutiva del diritto.

L'intelligenza artificiale, che non deve "competere", ma completare" le azioni umane, potrebbe, quindi, potenziare l'apertura dell'ordinamento giuridico al nuovo e responsabilizzare il giudicante, offrendogli gli strumenti che lo aiutano a motivare correttamente la decisione del caso concreto.

3.3 La funzione delle norme

Da quanto finora rilevato emerge che, nella prospettiva del tecnoumanesimo, è necessario rifuggire da atteggiamenti preconcetti di rifiuto oppure di entusiasmo fideistico nei confronti del contributo che l'intelligenza artificiale può fornire nella fase decisionale e promuovere una cooperazione virtuosa, che combini intelligenza umana e intelligenza artificiale e sia in grado di bilanciare i rispettivi pregi e difetti per tutelare i diritti della persona. La nuova tappa che si profila nel rapporto tra uomo e macchina, dunque, è rappresentata dalla complementarità e il processo costituisce un banco di prova per questa nuova collaborazione.

La demarcazione delle sfere di competenza, rispettivamente, dell'intelligenza artificiale e dell'intelligenza del giudice è una sfida che va affrontata, in primo luogo, attraverso una vasta campagna di alfabetizzazione digitale. Inoltre, le istituzioni pubbliche devono

[190] A. PUNZI, *Judge in the machine, e se fossero le macchine a restituirci l'umanità del giudicare?* cit, p.328.

[191] U. RUFFOLO, *Giustizia predittiva e machina sapiens quale "ausiliario" del giudice umano,* cit., pp.1 e ss.

elaborare un quadro normativo costituito da norme etiche e giuridiche che garantisca la tutela e la valorizzazione dell'essere umano.

Si muovono in questa direzione sia la *Carta etica europea sull'utilizzo dell'intelligenza artificiale nei sistemi giudiziari e negli ambiti connessi*, sia l'*AI Act* che, tuttavia, devono pongono problemi interpretativi.

Nella sfida dell'intelligenza artificiale, dunque, la Costituzione dovrà svolgere il ruolo di bussola, in modo che la disciplina europea e nazionale e le relative tutele siano sempre orientate a garantire la massima espansione delle garanzie dei diritti fondamentali.

In questo modo è possibile trasformare l'intelligenza artificiale, da «macchina che inquieta» a «macchina che sostiene, corregge e esalta» l'umanità.

CONCLUSIONI

Come stabilisce l'*AI Act* e ribadisce il d.d.l. n.1146 in discussione in Parlamento l'utilizzo di strumenti di IA può fornire sostegno al potere decisionale dei giudici o all'indipendenza del potere giudiziario, ma non deve sostituirlo: il processo decisionale finale deve rimanere un'attività a guida umana.

L'ancillarità dell'IA alla decisione del giudice dipende dalla natura della funzione giurisdizionale che non si può ridurre ad un calcolo matematico, né si può identificare in un sillogismo. Infatti, affinché si possa configurare un illecito da sanzionare, non è sufficiente l'accertamento della condotta ma bisogna anche esprimere un giudizio sulla colpevolezza del suo autore. L'imputazione psicologica richiede una valutazione cognitiva che solo la mente umana è al momento in grado di compiere.

La normativa europea, dunque, da una parte, afferma e ribadisce in modo netto il primato del giudice umano nella decisione, dall'altra, tuttavia, non può ignorare che i dati normativi, giurisprudenziali e dottrinali che lo stesso deve considerare nel suo lavoro, si sono moltiplicati. Il cervello umano, quindi, deve trarre beneficio dall'ausilio dell'intelligenza artificiale della macchina che è infinitamente più veloce ed esaustiva nel censire ogni dato utile e può dare informazioni e contributi "intelligenti" e, soprattutto, completi, anche nella fase dell'esame del fascicolo e della ricognizione di tutte le domande ed eccezioni delle parti.

Pertanto, la relazione giudice/IA è impostata non sul piano dell'equiordinazione, bensì sul diverso paradigma dell'ausiliarietà della seconda al primo.

La paura dell'IA è ingiustificata se è inserita nel contesto storico e filosofico corretto.

Da centinaia di anni, infatti, l'uomo costruisce versioni artificiali di sé stesso, dotandole di poteri superumani e progettandole per salvarsi dai propri limiti e condurre vite più sicure, più sane e più felici. Proprio perché sono così potenti, tali dispositivi possono diventare pericolosi. Tuttavia, essi entrano in un mondo dominato da Stati e imprese.

La chiave, quindi, è mantenere un discorso aperto e critico e, attraverso le norme giuridiche, assicurarsi che queste tecnologie servano il bene comune senza sacrificare i nostri diritti.

La disciplina dell'*AI Act* relativa all'uso dell'IA in ambito giudiziario rappresenta un primo passo in questa direzione e dimostra che l'UE intende promuovere la creazione di un unico ecosistema digitale centrato sull'individuo; un individuo che pretende dal

legislatore nuove situazioni giuridiche soggettive, i diritti digitali verso l'IA, ed è pronto ad assumersi i nuovi impegni individuali e solidali dettati dalla tecnica a guida umana.

BIBLIOGRAFIA

AA.VV., *Atti del Convegno "Nuovi scenari della giustizIA",* a cura di Maria Novella Campagnoli e Massimo Farina. Ottobre 2022, in *Rivista elettronica di Diritto, Economia, Management*, 2024, 1

ARDUINI S., La "scatola nera" della decisione giudiziaria: tra giudizio umano e giudizio algoritmico in *BioLaw Journal – Rivista di BioDiritto*, 2021, 2, pp.453-470

CARLEO A. (a cura di), La *decisione robotica*, Il Mulino, Bologna, 2019.

CASONATO C.,*Intelligenza artificiale e giustizia: potenzialità e rischi*, in *Diritto Pubblico Comparato ed Europeo (DPCE) on line*, www.dpceonline.it, 2020,3, pp.3369-3389

CAVALLO PERIN R., *Fondamento e cultura giuridica per la decisione algoritmica*, in U. SALANITRO (a cura di) Atti del Convegno "Smart la persona e l'infosfera", Università di Catania settembre/ottobre 2021, Pacini Giuridica, Pisa, 2022, p.89.

CORONA F., *La decisione del giudice tra precedente giudiziale e predizione artificiale,* in *Democrazia e Diritti Sociali*, 2023, 1, pp.83-97

CORRERA A., *Il ruolo dell'Intelligenza artificiale nel paradigma europeo dell'E-justice. Prime riflessioni alla luce dell'AI Act*, in *Quaderni AISDUE*, 2, 2024.

DE MINICO G., *Giustizia e intelligenza artificiale: un equilibrio mutevole*, in Rivista *Associazione Italiana Costituzionalisti (AIC)*, 2024, 2, pp.85-108.

LIPARI N., *Diritto, algoritmo, predittività*, in *Rivista trimestrale di diritto e procedura civile*, 2023, 3, pp.721-739

MORELLI A., *Il giudice robot e il legislatore naïf. La problematica applicazione delle nuove tecnologie all'esercizio delle funzioni pubbliche*, in *Liber amicorum per P.Costanzo*, 2020

ORTOLANI M.G., *La giustizia predittiva nell'ordinamento giuridico italiano e nei principali ordinamenti di common law*, in *Annali della Facoltà Giuridica dell'Università di Camerino*, 2023, 12, pp.1-23

PENASA S., *Intelligenza artificiale e giustizia: il delicato equilibrio tra affidabilità tecnologica e sostenibilità costituzionale in prospettiva comparata*, in *Diritto Pubblico Comparato ed Europeo* (DPCE) on line, www.dpceonline.it, 2022, 1, pp.297-310.

PUNZI A., *Il dialogo delle intelligenze tra umanesimo e tecnoscienza*, in *Persona e mercato*, 2023,2, pp. 161-168.

PUNZI A., *L'umanesimo digitale: verso un nuovo principio di responsabilità?* in *Democrazia e Diritti Sociali*, 2023, 1.

PUNZI A., *Mutamento di paradigmi o rottura antropologica? L'abito ermeneutico di Giuseppe Zaccaria e la giustizia digitale*, in *Rivista di filosofia del diritto*, 2023, 2, pp. 281-292

PUNZI A., *Decidere in dialogo con le macchine: la sfida della giurisprudenza Contemporanea*, in U. SALANITRO (a cura di) Atti del Convegno "Smart la persona e l'infosfera", Università di Catania settembre/ottobre 2021, Pacini Giuridica, Pisa, 2022, p. 261.

PUNZI A., *Difettività e giustizia aumentata. L'esperienza giuridica e la sfida dell'umanesimo digitale*, in *Ars interpretandi, Algoritmo ed esperienza giuridica*, 2021, p. 115.

ROMEO F., *Algoritmi giuridici di machine learning e controllo del giudicato interculturale*, in *Rivista di filosofia del diritto*, 2023, 1, pp.157 -168.

ROMEO F., *Giuseppe Longo, Matematica e senso. Per non divenire macchine*, in *La cultura*
2023, 2, pp.367 -372.

ROMEO F., *Algoritmi di giustizia ed equità nel diritto quando razionalità ed emozionalità convergono,* in *I-lex. Scienze Giuridiche, Scienze Cognitive e Intelligenza Artificiale*, www.i-lex.it, 2021, I.

ROMEO F., *Giustizia e predittività. Un percorso dal machine learning al concetto di diritto,* in
Rivista di filosofia del diritto, 2020, 1, pp. 107 – 124.

ROMEO F., *Esplorazioni nel diritto artificiale*, in *I-lex Scienze Giuridiche, Scienze Cognitive e Intelligenza Artificiale*, www.i-lex.it, 2004, I.

RUFFOLO U., *Machina iuris-dicere potest?,* in *BioLaw Journal – Rivista di BioDiritto*, 2021, 2, pp.399-410

SAPIENZA S., *Decisioni algoritmiche e diritto*, Giuffrè, Milano 2024.

SCIACCA M., *Algoritmo e giustizia alla ricerca di una mite predittività*, in Persona *e Mercato*, 2023, 1, p.1-15

TRIGGIANO A., *Gli algoritmi predittivi nel processo. Riflessioni storico giuridiche*, ESI, Napoli, 2023.